AN APPLICATION OF MULTIVARIATE DISCRIMINANT ANALYSIS AND CLASSIFICATION PROCEDURES TO RISK ASSESSMENT IN OPERATIONAL TESTING

A THESIS

Presented to

The Faculty of the Division of

Graduate Studies

By

Edward Dewey Simms, Jr.

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Operations Research

Georgia Institute of Technology

June 1975

AN APPLICATION OF MULTIVARIATE DISCRIMINANT ANALYSIS

AND CLASSIFICATION PROCEDURES TO RISK

ASSESSMENT IN OPERATIONAL TESTING

Approved:

Douglas C. Montgomery, Chairman

Russell G. Heikes

I. Gayden Thompson

Date approved by Chairman: June 10, 1975

ACKNOWLEDGMENTS

I gratefully acknowledge the aid, guidance, and prolonged assistance of Dr. Douglas C. Montgomery. His continual support and lack of restraint are indeed appreciated.

I would like to thank Drs. Russell G. Heikes and I. Gayden Thompson for serving on my reading committee. I am grateful for their suggestions and comments.

My appreciation is extended to Dr. Leslie G. Callahan and Mr. Thomas N. Burnette, Jr. for introducing and directing me into this area of study.

I wish to express my appreciation and very special thanks to my wife, Suzanne, for her continued interest and encouragement and for being so understanding during the time this paper was being prepared. In addition, I would like to thank Suzanne and Ms. Elaine Trammel for providing the typing expertise which produced the final rough drafts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SUMMARY

This research develops a methodology which determines an index used in the assessment of risk in Operational Testing. The risk assessment problem examined is that of preference statements regarding competing systems. In order to evaluate the competing systems, a multivariate statistical analysis of the systems is undertaken. Through the analysis of the multivariate distributions of each system and the overlap of these distributions, the index of risk is determined. Thus the index of risk is a measure of similarity of the competing systems.

In this thesis the probability of misclassification is used for the purpose of providing an index of similarity. Multivariate discriminant analysis and classification procedures are used so that the probability of misclassification can be determined. In keeping within the framework of operational testing, population distribution parameters are considered unknown, but sampling estimates are available.

In order to estimate the probability of misclassification, a computer simulation approach is advanced and incorporated into the proposed methodology. Through the generation of random observations from each population, use of linear and non-linear discriminant functions and optimal classification procedures, the probability of misclassification is determined.

It is envisaged that this methodology can effectively be utilized in the assessment of risk in the competing systems problem. While this methodology does not give an exact value to the risk, it does quantitatively determine an index of risk.

CHAPTER I

INTRODUCTION

Background

In the past few years, a great deal of attention has been focused on the Department of Defense's problems of unanticipated cost growth, system performance shortcomings, and failure to deliver equipment on schedule. The second of these problems has stirred additional interest in equipment testing prior to purchase. This desire to uncover system performance shortcomings has caused a vitalization of the entire testing procedure with much emphasis being placed on tests of the system in the actual user environment.

Effective 1 February 1975, the U.S. Army Operational Test and Evaluation Agency (OTEA) was given the mission to support the materiel acquisition processes by exercising responsibility for all operational testing [8]. Some of the functions of OTEA under this mission are:

1. Insure that user testing is effectively planned, conducted, and evaluated with emphasis on adequacy, quality, and credibility of all user testing.
2. Develop and promulgate user test and evaluation methodology.
3. Develop measures of effectiveness, necessary to detect difference in the military utility, operational effectiveness, and operational suitability.
4. Provide user test input for Army submissions to Office of the Secretary of Defense and Congress.

An operational test is that test and evaluation conducted to estimate the prospective systems's utility, operational effectiveness, and operational suitability. This test is accomplished with typical user operators, crews, or units in as realistic an operational environment as possible. One of the major goals of OT is to provide data to estimate the system's desirability, considering systems already available.

In the early part of the operational test sequence, the test may involve several competing systems. In OTI, the test must be able to evaluate the relative merits of available competing prototypes/systems from the aspect of military utility. OT II may also be characterized by comparative testing between competing systems and the item being replaced. Hence one of the objectives of operational testing and OTEA is an independent evaluation of competing systems resulting in some statement of relative attributes and preference.

The major obstacles in achieving the above objectives lie in the uncertainties inherent in testing. Tests run on a prototype in the hands of a relatively few user/operators may not accurately depict the true characteristics of the tested system. Thus the results of an OT evaluating competing systems may not correctly establish the relative qualities of each system.

In order to delineate and evaluate these uncertainties in testing, the Department of Defense has inaugurated major programs dealing with risk analysis. Formal recognition and emphasis on risk analysis resulted from a July 31, 1969 memorandum in which Secretary Packard cited inadequate identification and consideration of risks in major programs as a problem area requiring action [50]. Hence there has been much effort expended in the conduct of a risk analysis on new equipment.

Risk Analysis

There has been a great deal written about risk analysis as it is applied to the system acquisition process. But mainly because risk analysis is an art and an uncertainty; it is extremely difficult to standardize. In fact, the Air Force Academy Risk Analysis Study states,

> . . . in the context of the system acquisition process, it (Risk Analysis) is so nebulous and ill-defined that identifying basic concepts, developing meaningful guidelines, and describing effective methodologies are, in themselves, difficulties of the first magnitude. [50]

Even though this statement is foreboding, there have been considerable advances in the methodology of risk analysis. In large systems the analysis consists of the study of such things as the probability of cost overruns, probability of failure to meet development timetables and probability of performance shortcomings. Risk analysis can be viewed as the process of combining the risk assessment with alternate courses of action in an iterative cycle. As the testing environment becomes constrained, new alternatives must be generated and the appropriate risk assessed. Therefore, risk analysis is a highly coordinated examination of all factors which affect the risk of acquiring systems.

The coordination which is necessary is between two different areas. First, there are the various strategies or alternatives which are available for dealing with problems which arise from uncertainty. In conducting a risk analysis, OTEA is required to generate an initial set of alternatives, assess the risk in these, modify the alternatives to reduce the risk, reassess the risk, and continue this process until one alternative may be chosen as the optimum end. These alternatives are strategies that are feasible in the constrained environment. In the case of the competing systems problem, the alternative strategies may be the number of samples taken or the number of attributes examined in the operational test. As the constraints on funds and time are changed, the alternatives change necessitating a new risk analysis.

The second area of coordination in the analysis is the estimation (either quantitative or qualitative) of what the risk associated with a specific course of action actually is--risk assessment.

Risk Assessment

OTEA has as one of its implied missions, the responsibility of conducting risk analyses on its evaluation of candidates. Given the wide spectrum of systems that must be evaluated by OTEA, no one method of analysis will always be optimum. But all methods of analysis will always require the identification and assessment of risk. As defined by the Air Force Academy Risk Analysis Study Team, risk assessment is, "a comprehensive and structured process for estimating the risk associated with a particular alternative course of action."

Risk assessment is not a new notion. The problem of performance uncertainty in operational testing has been addressed previously. The process, however, was largely intuitive, incomplete, and informal. It was intuititive in that a structured quantitative approach often gave way to intuition. It was incomplete in that detailed analyses of isolated aspects of the problem were rarely brought together and integrated in a broader analysis. And it was informal in that the results of the assessment were often not written and explicitly incorporated into the review (approval) control process.

In the case of the competing systems evaluation, one such measure of risk may be the probability of making an incorrect preference statement regarding the systems. Since the true operational characteristics of each are not known with certainty, there is a finite probability that the preference statement regarding these characteristics offered by OTEA may be incorrect.

Procedure

At the present time, operational characteristics are analyzed and variables pertaining to these characteristics are defined. These variables represent measures that when determined, effect the evaluation of the operational characteristics. After the variables have been defined, the operational test is developed and executed, resulting in data on these variables for each of the competing systems. One method of comparison is to evaluate each variable separately and compare the candidates. Associated with each comparison would be some finite risk, therefore, at completion of OT, there would be many comparative evaluations with their associated risks. The problem arises as to how to combine these risks when the preference statement is made.

One method of combining the comparisons on all of the variables into one comparison is through the use of multivariate statistical analysis. Tatsuoka [48] describes multivariate statistical analysis as "that branch of statistics which is devoted to the study of multivariate (or multidimensional) distributions and samples from those distributions." For the applied statistician and researcher who uses statistics as a tool, this definition would not be adequate. Press

[39] gives a more applied characterization by stating that multivariate analysis is "that branch of statistics that is devoted to the study of random variables which are correlated to one another." If some type of correlation can be determined, then the behavior of one provides some knowledge about the behavior of the other. Therefore, through the use of multivariate statistics, the total risk can be evaluated.

In order for this risk to be evaluated, an index must be developed that will reflect the risk involved. One method of indexing the risk is to examine the question, "Is there a difference between the prototypes and how much is it?" The index used in this thesis is the probability of misclassifying an observation from one candidate population into another candidate population. Overall and Kleit wrote, "Calculation of the proportion of misclassification provides a meaningful index of the degree of separation between the two groups" [38]. If the probability is small; then, the risk involved in making a statement regarding the differences between the candidates is small. Likewise, if the probability of misclassification is high, then the risk involved in stating that the systems are operationally different is high. Thus the objective of the methodology presented is to calculate the probabilities of misclassifying observations from one population into another population.

Objective of Research

The primary objective of this research is to develop a methodology with which to assess the risk involved in competing systems in operational testing. Risk, as used in the context of this thesis, is defined as the probability that a preference statement based on the

OT evaluation of competing systems is incorrect. Therefore, the purpose of this research is the development of a methodology that will quantitatively assess the probability of an incorrect preference statement in an operational test environment.

The accomplishment of this objective entails a study of discriminant functions and classification procedures in multivariate statistics. Both linear and nonlinear discriminant functions are examined. Computer simulation techniques also are examined with special emphasis placed on variate generation techniques. An example problem using data from the Squad Automatic Weapons Test is used in Chapter III to clarify the methodology presented.

CHAPTER II

METHODOLOGY

Introduction

The methodology proposed to assess the risk index of probability of misclassification can be partitioned into three major areas: discrimination, classification, and data analysis.

The first major area is the determination of the discriminant function to be used. Two major types of discriminant functions are considered and evaluated. The linear discriminant function has as its objective the collapsing of multivariate statistical distribution to a univariate statistical distribution. Once the transformation from the multivariate space to the univariate space has been made, standard statistical operations can be carried out. One of the major advantages of the linear discriminant function is that by its mathematical nature, inferences on the relative importance of each variate regarding discrimination can be made. A major disadvantage is the assumption of common dispersion of the candidate populations.

The second type of discriminate function is the nonlinear discriminant function. This type of discrimination examines the multivariate distribution and discriminates according to densities. An advantage of this discriminant function is that common dispersion is not assumed, but inferences on the individual variates are not possible.

The next major area of the methodology is classification. If all the parameters for each of the populations were known, the probability of misclassification would have a closed form solution. When estimates for the parameters are used, the probability of misclassification must be estimated. Several methods of estimation are examined and a frequency interpretation of the probability of misclassification using computer simulation is suggested. This frequency of misclassification is based on the classification rule that incorporates both cost and prior information.

Finally, the data from the sample of candidate populations must meet certain restricting assumptions. Methods for examining the data and transformation of the data to a usable form are discussed in the Data Analysis section.

Linear Discrimination

One of the first discussed and most widely used forms of discrimination is linear or simple discriminant analysis. In linear discrimination, the objective remains the same in that an index is sought that will reflect the information contained in the multiple measurements. As the name implies, the index obtained in using linear discrimination is a weighted linear combination of the variates in the observation vector. By assigning appropriate weighting coefficients, several scores (values of variates) can be transformed to a single score (value of the index) which has maximum potential for distinguishing between members of two groups. In this manner, the multivariate problem is actually reduced to a simple univariate problem, and assignment of individual observations between two groups depends upon the value of a single variable.

Perhaps a geometric interpretation of linear discriminant analysis would be helpful at this time. In Figure 2-1 two populations are represented by measurements on each of the variables (x_1 and x_2). It is assumed that the scores for each group are distributed as bivariate normal with the same dispersion matrices. Therefore, the two ellipses represent the same constant density for the two bivariate normal distributions. Notice that there is some overlap between these ellipses. Clearly, the closer the mean vectors, the greater will be the overlap for any constant density and the more difficult it will be to discriminate between the populations.

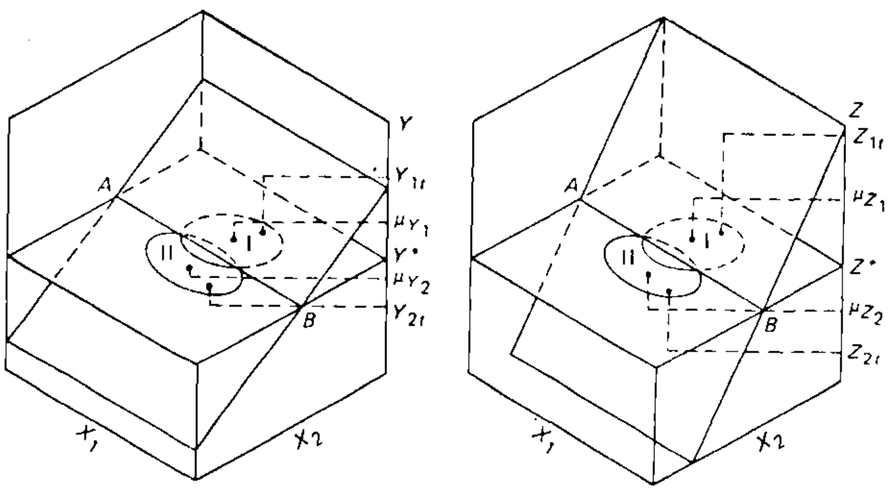

Figure 2-1. Constant Density Ellipses, Discriminant Planes, and Projections [9, p. 231]

The procedure in linear discriminant analysis is to find a linear combination of the measures (x_1 and x_2) such that the distributions for the two groups will possess "little" overlap. The linear function

$$Y_{it} = \beta_1 X_{i1t} + \beta_2 X_{i2t} \qquad \begin{array}{l} i = 1,2 \\ t = 1,2,\ldots,n \end{array}$$

is called a linear discriminant function with unknown coefficients $\{\beta_i\}$. The subscript i represents the group, and the subscript t refers to the observation number within a group. Geometrically, the above equation defines a plane. The projection of X_{i1t} and X_{i2t} on the plane transforms the two-dimensional observations into a one-dimensional discriminant score, Y_{it}. The plane cuts the ellipses along the line AB which passes through their points of intersection, and the projection of line AB is y*. Thus the plane cuts the isodensity ellipses with most of ellipse I being under the plane and most of ellipse II being above the plane. Observation vectors which project onto the y axis above y* are classified into Group I. Similarly, vectors which project onto the y axis below y* are classified into Group II. The mean vector for Group I and II will project onto the y axis as μ_{y1} and μ_{y2}. Misclassification occurs whenever an observation vector from Group I projects below y* and vice versa. From the figure the observations that will be misclassified (the area of misclassification) lie in the overlap of the two isodensity ellipses. If the discriminant plane passes through the points of intersection of the two isodensity ellipses, the overlap between the two groups will be minimized [14]. However, there are an infinite number of planes that pass through the isodensity ellipse intersections such as plane Z in the right figure.

In order to evaluate the desirability of each plane, let $(\mu_{y1} - \mu_{y2})^2$ and $(\mu_{z1} - \mu_{z2})^2$ represent the separation of the two groups by the respective planes. Each of these distances measure the variation <u>between</u> the means of the discriminant scores for each group. In this respect, the plane which results in the largest distance between groups would be superior since separation of the groups is desired. On the other hand, small variation within the projection, $(Y_{1t} - \mu_{y1})^2$ and $(Z_{1t} - \mu_{z1})^2$ is desired because large within-group variation tends to negate the significance in a statistical sense of the distance between the projected means. Thus the optimal discriminating plane is the plane which maximizes the following ratio:

$$\lambda = \frac{\text{between-group variation}}{\text{within-group variation}}$$

so that the between-group differences will be large relative to within-group scatter.

For two populations, the projections of the mean vectors are given by

$$\mu_{y_1} = \beta_1 \mu_{x_{11}} + \beta_2 \mu_{x_{12}} + \dots \beta_p \mu_{x_{1p}}$$

and

$$\mu_{y_2} = \beta_1 \mu_{x_{21}} + \beta_2 \mu_{x_{22}} + \dots \beta_p \mu_{x_{2p}}$$

Thus, under the assumption of common dispersion, the within-groups sum of squares is

$$\sum_{i=1}^{2} \sum_{t=1}^{n_i} (y_{it} - \mu_{yi})^2 = \sum_{i=1}^{2} \sum_{t=1}^{n_i} [\sum_{k=1}^{p} \beta_k (x_{ikt} - \mu_{yik})]^2$$

$$= \sum_{j=1}^{2} \sum_{t=1}^{n_i} [\beta_1 (x_{j1t} - \mu_{yj1}) + \cdots + \beta_p (x_{ipt} - \mu_{xip})]^2$$

Introducing vector notation,

$$\underset{\sim}{\beta}' = [\beta_1, \beta_2, \ldots, \beta_p]$$

$$\underset{\sim}{\mu}_i' = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{ip}]$$

$$\underset{\sim}{d} = \underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2$$

and define

$$\underset{\approx}{X}_i = E\left\{\begin{bmatrix} x_{i11} & x_{i21} & \cdots & x_{ip1} \\ x_{i12} & x_{i22} & & \vdots \\ \vdots & & & \\ x_{i1n} & & & x_{ipn} \end{bmatrix} - \begin{bmatrix} \mu_{i1} & \mu_{i2} & \cdots & \mu_{ip} \\ \mu_{i1} & & & \\ \vdots & & & \\ \mu_{i1} & & & \mu_{ip} \end{bmatrix}\right\}$$

Then

$$\sum_{i=1}^{2} \sum_{t=1}^{n_i} (y_{it} - \mu_{yi})^2 = \underset{\sim}{\beta}' \underset{\approx}{X}_1' \underset{\approx}{X}_1 \underset{\sim}{\beta} + \underset{\sim}{\beta}' \underset{\approx}{X}_2' \underset{\approx}{X}_2 \underset{\sim}{\beta}$$

$$= \underset{\sim}{\beta}' (\underset{\approx}{X}_1' \underset{\approx}{X}_1 + \underset{\approx}{X}_2' \underset{\approx}{X}_2) \underset{\sim}{\beta}$$

However the pooled covariance matrix is defined as:

$$\underset{\sim}{\Sigma} = \frac{1}{n_1 + n_2} (\underset{\sim}{X}_1' \underset{\sim}{X}_1 + \underset{\sim}{X}_2' \underset{\sim}{X}_2)$$

Thus the within-group variation is

$$\underset{\sim}{\beta}'[(n_1 + n_2)\underset{\sim}{\Sigma}]\underset{\sim}{\beta}$$

The between-group variation can be expressed as the square of the difference of the projected population means.

$$(\mu_{y_1} - \mu_{y_2})^2 = \underset{\sim}{\beta}'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)'\underset{\sim}{\beta}$$

$$= \underset{\sim}{\beta}' \underset{\sim}{d} \underset{\sim}{d}' \underset{\sim}{\beta}$$

Thus the function to be maximized is

$$f(\beta_1, \beta_2, \ldots, \beta_p) = \frac{\underset{\sim}{\beta}' \underset{\sim}{d} \underset{\sim}{d}' \underset{\sim}{\beta}}{\underset{\sim}{\beta}'[(n_1 + n_2)\underset{\sim}{\Sigma}]\underset{\sim}{\beta}}$$

Taking partial derivative with respect to $\underset{\sim}{\beta}$ we have

$$\frac{\partial f(\beta_1, \beta_2, \ldots, \beta_p)}{\partial \beta} = C \underset{\sim}{d} - \underset{\sim}{\Sigma} \underset{\sim}{\beta} = 0 ,$$

where C is a nonzero constant. Hence the weighting coefficients are given by $\underset{\sim}{\beta} = \underset{\sim}{\Sigma}^{-1} C \underset{\sim}{d}$. Since C is a constant and the $\underset{\sim}{\beta}$ weights are applied to both populations, without loss of generality C can be set equal to one.

Thus

$$\underset{\sim}{\beta} = \underset{\sim}{\Sigma}^{-1} \underset{\sim}{d} ,$$

which is equivalent to solving simultaneously the following set of p equations with unknowns.

$$\begin{aligned} \beta_1\sigma_{11} + \beta_2\sigma_{12} \cdot \cdot \cdot \beta_p\sigma_{1p} &= d_1 \\ &\vdots \\ \beta_1\sigma_{p1} + \quad \cdot \cdot \cdot \quad \beta_p\sigma_{pp} &= d_p , \end{aligned}$$

where the σ's are elements of the within-groups variance-covariance matrix [33].

It must be noted that by substituting this solution for $\underset{\sim}{\beta}$ into $f(\beta_1,\beta_2,\ldots,\beta_p)$ the Mahalanobis D^2 or generalized distance is attained.

$$\Delta^2 = \underset{\sim}{d}' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{d}$$

Thus a vector of weights $\underset{\sim}{\beta}$ has been found which maximizes the Mahalanobis distance between the two populations.

In seeking to interpret the discriminant function, it is desired to know which of the original p variables contributes most to the function. In the past, investigators have attempted to define or describe the nature of the discriminant function by examining relative magnitudes of the weighting coefficients. For this purpose, comparison of the relative magnitudes of the combining weights is inappropriate because these are weights to be applied to the variates in raw-score scales, and are hence affected by the particular unit used for each variable. To eliminate the spurious effects of units on the magnitudes of combining weights, the comparison must be made of the weights as they would be applied to the variates in standardized form. A method of assessing these standardized weights is multiplying each original measure weight by the standard deviation of the corresponding variable as computed from the pooled-within groups covariance matrix. This amounts to multiplying each element in $\underset{\sim}{\beta}$ by the square root of the corresponding diagonal element of $\underset{\sim}{\Sigma}$. Thus

$$\beta_i^* = \sqrt{\sigma_{ii}}\ \beta_i \qquad i = 1,2,\ldots,p$$

are defined as the standardized discriminant weights. The relative contribution of the ith variable to the discriminant function can be evaluated by the magnitude of β_i^* in comparison with the other weights β_j^*.

By using the discriminant weighting coefficients $\underset{\sim}{\beta}$, the multiple measurement vector has been transformed into a univariate index under the assumption of multivariate normality and common dispersion for the two populations. The index variable for each population will be normally distributed with mean

$$\mu_{y_i} = \sum_{j=1}^{p} \beta_j\ \mu_{ij}$$

and variance

$$\sigma_y^2 = \sum_{\substack{i=1 \\ i \neq j}}^{p} \sum_{j=1}^{p} \beta_i\ \sigma_{ij}\ \beta_j + \sum_{i=1}^{p} \beta_i^2\ \sigma_{ii}^2$$

This means that the deviation of an individual discriminant-function score from each of the group means can be regarded as a unit-normal deviate or Z score

$$Z_{y_i} = \frac{y - \mu_{y_i}}{\sqrt{\sigma^2_y}}$$

Thus for any particular discriminant-function score, say y_c, the Z-score deviation from each group mean can be computed. The unit normal distribution can be used to obtain an estimate of the probability of deviation from each group mean as large as that represented by any particular Z-score value. A particular discriminant-function score value y_c can be used as a decision point for classifying observations into two groups. The proportion of misclassifications can be read from the unit-normal distribution tables after transforming y_c to Z-score from using the above equation. By varying the location of the decision point, the proportion of misclassification with the two groups can be changed.

Figure 2-2 shows a particular discriminant-function value falling between the two group means μ_{y1} and μ_{y2}. If every observation having a discriminant-function value (y_i) less than y_c was classified into group II, the proportion of observations actually belonging to group I would be represented by the area under G_1 to the left of y_c. To determine this

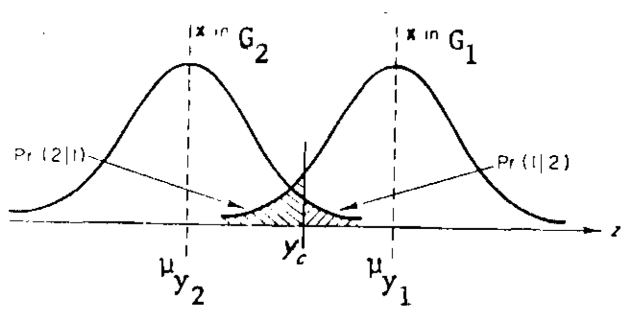

Figure 2-2. Distributions of Z-Scores [1, p. 237]

proportion, transform y_c to Z-score form

$$Z_y = \frac{y_c - \mu_{y1}}{\sqrt{\sigma_y^2}}$$

and then look up the area in the smaller portion of the unit-normal curve corresponding to Z_y.

Anderson (1958) showed that if the prior probability than an observation comes from the two groups is equal, then y_c will lie midway between μ_{y1} and μ_{y2}. Thus under the appropriate assumptions

$$Pr(2|1) = Pr(1|2) = \Phi\left(\frac{\frac{|\mu_{y1} - \mu_{y2}|}{2}}{\sigma_y}\right)$$

where $Pr(2|1)$ is the probability of classifying one observation from group 1 into group 2.

It may not always be the best solution to choose a cutting point for classification which equates the probabilities of error in assignment of individuals between two groups. The relative numbers of observations expected to belong to the two populations may be an important consideration, since the actual numbers of observations misclassified will be equal to the probabilities of misclassification times the relative numbers in the two populations. Thus the minimization of

$$P_\varepsilon = q_1 P_\varepsilon^{(1)} + q_2 P_\varepsilon^{(2)}$$

where P_ε = expected probability of misclassifying an observation and $P_\varepsilon^{(i)}$ = probability of misclassifying an observation from group (i) will reduce the expected number of observations misclassified. In order to minimize P_ε the cutting point y_c will have to be adjusted.

Rao (1965) showed that under the assumption of multivariate normality and common dispersion, the optimal cutoff point is

$$y_c = \frac{\mu_{y1} + \mu_{y2}}{2} + \ln\left(\frac{q_2}{q_1}\right)$$

If $q_1 = q_2 = 1/2$ then the cutoff point y_c will be again equal to

$$y_c = \frac{\mu_{y1} + \mu_{y2}}{2}$$

A further refinement of the linear discriminant procedure is to incorporate the concepts of costs of misclassification. This involves introducing a quantity $C(i|j)$ which is the cost or loss due to classifying an observation from population j into population i. The probabilities of error that one may be willing to accept will depend upon the relative seriousness of misclassification of observations from the two populations, which may be quite different in the two cases. The driving force behind this refinement is the minimization of the expected cost of misclassification. Thus a cutting point would be chosen that would minimize

$$TC = q_1\ C(2|1)\ P_\varepsilon^{(1)} + q_2\ C(1|2)\ P_\varepsilon^{(2)}$$

where TC = expected cost of misclassification, and

$C(i|j)$ = cost of classifying an observation from j into i.

Again referring to Rao and under the assumptions of multivariate normality and common dispersion, the optimal cutoff point is

$$y_c = \frac{\mu_{y1} + \mu_{y2}}{2} + \ln \frac{q_2 C(1|2)}{q_1 C(2|1)}$$

If the prior probabilities are equal and the costs of misclassification are equal, then the optimal cutoff point will remain the midpoint between the two univariate classification indices.

Nonlinear Discrimination

The classification problem as delineated earlier amounts to seeking an answer to the question: "Which of these two populations does this observation 'resemble' the most, in terms of a specified set of measurable characteristics?" That is, there are two well-defined populations with p variables that are known or are deemed to be important in differentiating among the two populations. Subsequently, an observation is encountered whose population membership is unknown, but for whom measures on these same p variables are available. The objective of the classification process is to classify this observation as a member of one or the other of the two populations--the one with which the observation shows greatest "resemblance" in terms of these p variables.

The crux of the matter lies in how to define "resemblance" in this context. Various measures of profile (or pattern) similarity and of distance (that is, dissimilarity) have been proposed in the literature [Mahalanobis, 1936; Cattell, 1944; DaMas, 1949: Crombach and Gleser, 1952].

The classification procedure to be presented assumes a multivariate normal distribution for the vector variable in each of the populations. The notion of "swarm" is used for the plot in the measurement space of points representing all the members of a single population, each point being located by treating the member's vector of measurement scores as coordinates of a single point in p-dimensional space [14]. A multivariate normal swarm is very dense in the region of the population centroid and thins out in all directions away from the centroid.

The swarm may be elongated in some directions as a function of covariances among the measurements and thus the rate of thinning in any direction is a resultant of variances and covariances. The normal swarm is hyper-ellipsoidal, meaning loosely that the projection of the swarm on any plane passing through the centroid is elliptical. This classification procedure conceives of a boundary in the swarm within which a proportion of the observations will be found when each observation is represented by the deviation of its score vector from the centroid. Each such boundary is one out of a set of concentrically nested ellipsoids. Such a boundary is the locus of all points for which a generalized distance function from the centroid is a constant value.

The familiar chi-squared or χ^2 statistic is used to serve as a measure of dissimilarity. That is

$$\chi^2 = \underset{\sim}{\underline{X}}' \, \underset{\sim}{\Sigma}^{-1} \, \underset{\sim}{\underline{X}} \, ,$$

where

$$\underset{\sim}{\underline{X}} = [x_1-\mu_1, x_2-\mu_2, \ldots, x_p-\mu_p]$$

$$\underset{\sim}{\Sigma} = \begin{bmatrix} \sigma_1^2 & p\sigma_1\sigma_2 & \cdots & p\sigma_1\sigma_p \\ \vdots & \cdots & & \cdots \end{bmatrix}$$

The χ^2 statistic is a reasonable choice, since the larger the χ^2 value of an observation with reference to a given population, the farther away (in the generalized distance sense) is the point $(x_1, x_2, \ldots, x_{pi})$ representing the observation's set of scores from the centroid $(\mu_{1i}, \mu_{2i}, \ldots, \mu_{pi})$. Hence the observation may be said to be the more deviant from the "average member" of that population, the larger its χ^2 value. Conversely, an observation with a small χ^2 value with reference to a group is "closer" to the average member of that group, and may be said to resemble that group. Furthermore, since the reference population is adequately described by means of a multivariate normal distribution of the p variables, then knowledge of an observation's χ^2 value allows the estimation of the percentage of observations in the group that are "closer to" or "farther from" the group centroid than is that observation. This is because the χ^2 value determines the particular centile ellipsoid on which a given point lies. This derives from the fact that a p-variate normal population $N(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$ has density

$$\phi(X_1, X_2, \ldots, X_p) = (2\pi)^{-p/2} \, |\Sigma|^{-1/2} \, \exp(-\chi^2/2)$$

Thus a simple classification scheme, which may be called the minimum chi-square rule, would be as follows:

> Compute the χ^2 value of the unclassified observation with respect to each of the two populations, and assign the observation to that population with respect to which its χ^2 value is the smallest.

This rule has the property of minimizing the probability of misclassification when the two populations have multivariate normal distributions with equal dispersion matrices.

Actually, the objective of the procedure is to find the probability that observation i is a member of the jth population. The hypotheses regarding the group membership of the observation when two populations are under study and the ith subject must be classified into one and only one population is:

$$Pr(H_j|\underset{\sim}{X}_i) \text{ , } i = 1,\ldots,N \quad \text{and} \quad j = 1,2$$

which reads: The probability of hypotheses j given the score vector $\underset{\sim}{X}_i$. Hypothesis j states that the observation is a member of population j. There are two such hypotheses evaluated for each subject and the maximum likelihood classification rule is to assign i to group j if $Pr(H_j|X_i)$ is largest [14].

The relation of the probability of the hypothesis that i belongs to group j to the cumulative probability for the distance function, χ^2_{ji}, is:

$$Pr(H_j|\underset{\sim}{X}_i) = 1 - P(\chi^2_{ij})$$

By this method the probability of group membership is simply the inverse cumulative function of the generalized distance tabled as chi-square with p degrees of freedom, where p is the dimension of the classification space.

Roulon (1967) and his associates have named this method of evaluating group membership the "centour" method. They define a centour score as:

$$100[1 - P(\chi^2_{ij})]$$

That is, the centour for observation i in population j is 100 times the probability of obtaining a larger value of χ^2_{ij}.

Since by the centour classification procedure the probability of membership in each population is inversely monotonic with the chi-square value for each group, it is equivalent to the minimum chi-square classification rule. Under the proper assumptions of multivariate normality and common dispersion matrices and with equal prior probabilities of membership, the optimal boundary will be a hyperplane between the two centroids with the probability of misclassification being equal for both populations.

One problem with this rule is that if the dispersion matrices of the two populations are not equal, overassignment of observations can occur. If $|\Sigma_j| > |\Sigma_k|$, group j will tend to be overassigned because a given centour for group j will enclose a larger region of the classification space than will the same centour for group K.

In order to dismiss the assumption of equal dispersion the density function for the multivariate normal is used:

$$P_{jk} = (2\pi)^{-P/2} \; |\Sigma_j|^{-1/2} \; \exp(- \frac{\chi^2_{ji}}{2})$$

where

$$\chi^2_{ji} = (\underset{\sim}{X}_i - \underset{\sim}{\mu}_j)' \; \underset{\sim}{\Sigma}_j^{-1} \; (\underset{\sim}{X}_i - \underset{\sim}{\mu}_j)$$

and P_{ji} = probability of X_i given population j.

Using Bayes Theorem:

$$\Pr(H_j|\underset{\sim}{X}_i) = \frac{\Pr(\underset{\sim}{X}_i|H_j)}{\Pr(\underset{\sim}{X}_i|H_1) + \Pr(\underset{\sim}{X}_i|H_2)}$$

and

$$\sum_{j=1}^{2} \Pr(H_j|\underset{\sim}{X}_i) = 1.0$$

Thus the classification rule is

$$H_j \quad \text{if} \quad \Pr(H_j|\underset{\sim}{X}_i) > \Pr(H_k|\underset{\sim}{X}_i) \qquad \begin{matrix} k = 1,2 \\ k \neq j \end{matrix}$$

This classification rule allows the discriminant boundaries to be a nonlinear function (Figure 2-3).

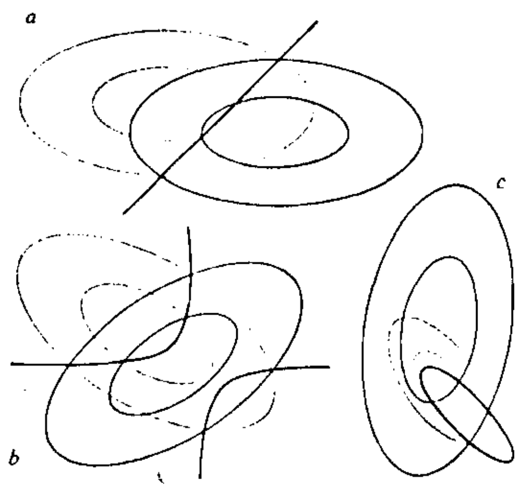

Figure 2-3. Intersections of Corresponding Equiprobability Contours. (In (a) the variances and the covariance are equal for both distributions, and the discriminant curve becomes a straight line. [51, p. 266])

This classification criterion can be amended for prior information on population membership in much the same way as the linear discriminant function. P_{ji} can be easily redefined to reflect these prior probabilities as

$$P_{ji} = q_j \; |\Sigma_j|^{-1/2} \exp \left(\frac{-X_{ji}^2}{2} \right)$$

where q_j is the prior probability that an observation comes from population j.

In like manner the costs of misclassification can be included in this approach. As in the linear case, the objective of the entire procedure is to minimize the total expected costs of misclassification. The classification rule according to density may be rewritten as:

$$H_j \quad \text{if} \quad \frac{\Pr(H_j|\underset{\sim}{X}_i)}{\Pr(H_k|\underset{\sim}{X}_i)} \geq 1 \qquad \begin{array}{l} k = 1,2 \\ k \neq j \end{array}$$

using prior information the rule is

$$H_j \quad \text{if} \quad \frac{q_j(H_j|\underset{\sim}{X}_i)}{q_k(H_k|\underset{\sim}{X}_i)} \geq 1 \qquad \begin{array}{l} k = 1,2 \\ k \neq j \end{array}$$

Considering the costs of misclassification:

$$H_j \quad \text{if} \quad \frac{C(k|j) \; q_j \; \Pr(H_j|\underset{\sim}{X}_i)}{C(j|k) \; q_k \; \Pr(H_k|\underset{\sim}{X}_i)} \geq 1 \qquad \begin{array}{l} k = 1,2 \\ k \neq j \end{array}$$

where $C(k|j)$ = cost of classifying an observation from population j into population k.

Thus this procedure minimizes the expected cost of misclassification given by

$$\sum_{j=1}^{2} q_j \{\sum_{\substack{j=1 \\ i\neq j}}^{2} C(i|j)\ Pr(i|j)\},$$

where $Pr(i|j)$ = probability of classifying an observation from population j into population i.

Disregarding the costs of misclassification, that is assuming the $C(i|j)$ equal, this procedure minimizes the expected probability of misclassification given by

$$\sum_{j=1}^{2} q_j \{\sum_{\substack{j=1 \\ i\neq j}}^{2} Pr(i|j)\}$$

Sample Estimation

The problem with the closed-form solution to the probability of misclassification problem is that the population parameters are usually estimated by sample observations from the populations. Thus the mean vectors μ_1 and μ_2 and the covariance matrix $\underset{\sim}{\Sigma}$ are unknown. If $x_{11},\ldots,x_{1n_1}$ and $x_{21},\ldots,x_{2n_2}$ are independent random samples from population 1 and population 2 respectively, then the mean vectors μ_i can be estimated by the sample mean vectors $\underset{\sim}{X}_i = (\overline{x}_{i11}\ldots\overline{x}_{ip})$ and $\underset{\sim}{\Sigma}$ by the pooled covariance matrix $\underset{\sim}{S}_*$.

Hence, the mean variable j in population i is estimated by

$$\overline{x}_{ij} = \frac{\sum_{t=1}^{n_i} x_{ij_t}}{n_i} \qquad \begin{array}{l} i = 1,2 \\ j = 1,2,\ldots,p \end{array}$$

The population mean vector $\underset{\sim}{\mu}_i$ is estimated by

$$\underset{\sim}{\overline{X}}_i = [\overline{x}_{i1},\ldots,\overline{x}_{ip}]$$

The population mean of the univariate index μ_{yi} is estimated by

$$\overline{Y}_i = \frac{\sum_{t=1}^{n_i} y_{it}}{n_i} \qquad i = 1,2$$

Thus

$$\underset{\approx}{X}_i = \begin{bmatrix} x_{i11} & x_{i21} & \cdots & x_{ip1} \\ x_{i12} & & & \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ x_{i1n} & & \cdots & x_{ipn} \end{bmatrix} - \begin{bmatrix} \overline{x}_{i1} & \cdots & \overline{x}_{ip} \\ \overline{x}_{i1} & & \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \overline{x}_{i1} & \cdots & \overline{x}_{ip} \end{bmatrix}$$

The common dispersion matrix $\underset{\sim}{\Sigma}$ can be estimated by the pooled covariance matrix

$$\underset{\sim}{S}_* = \frac{1}{n_1+n_2-2} (\underset{\approx}{X}_1' \underset{\approx}{X}_1 + \underset{\approx}{X}_2' \underset{\approx}{X}_2)$$

In such a situation, it is not possible to derive a classification procedure which is optimal (in the sense of minimizing expected cost of misclassification) as was done earlier. However, it is shown by Anderson (1958) in Theorem 6.5.1 that if consistent estimates are substituted for the parameters in the generalized procedure then the resulting expected cost of misclassification becomes minimized as n_1 and $n_2 \to \infty$ with constant ratio.

In the linear discriminant function the weighting coefficients will be determined by,

$$\underset{\sim}{\beta} = \underset{\sim}{S}_*^{-1} \underset{\sim}{d}$$

where

$$\underset{\sim}{d} = (\overline{\underset{\sim}{X}}_1 - \overline{\underset{\sim}{X}}_2)$$

Under the procedure of the nonlinear discriminant function due to Welch, the likelihood ratio criterion leads to the criterion: assign observation to group 1 if

$$f_1/f_2 > 1$$

where f_1 and f_2 represents the respective densities and under the assumptions of equal prior probabilities and equal costs of misclassification.

The objective of both methods of discrimination is to minimize the expected costs of misclassification by minimizing the probability of misclassification. Letting $D_s(x)$ represent the value of the discriminant function, the probabilities of misclassification are:

$$P_L = P(D_s(x) < K | G_L, \overline{\underset{\sim}{X}}_1, \overline{\underset{\sim}{X}}_2, \underset{\sim}{S}_*)$$

$$P_M = P(D_s(x) > K | G_M, \overline{\underset{\sim}{X}}_1, \overline{\underset{\sim}{X}}_2, \underset{\sim}{S}_*) \ .$$

where K is the cutting point in linear discriminant analysis or 1 in Welch's (1939) method [53].

The exact distribution of $D_s(x)$ is quite complicated. It has been studied by Wald (1944), Anderson (1958), Sitgraves (1952), Kabe (1963), and Okamoto (1963) among others. If the parameters of the distribution are known, the problem reduces to $P_1 = P_2 = \Phi(-\Delta/2)$ where $\Delta^2 = (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)' \, \Sigma^{-1}(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$ is Mahalanobis' distance and Φ is the cumulative normal distribution. This represents a limiting factor which cannot be improved upon, and unfortunately, cannot be found.

Currently there are several methods in use that estimate the probability of misclassification of sample discriminant functions. The techniques may be divided into two classes: those using a sample to evaluate the discriminant function, and those using properties of the normal distribution. The former may be considered empirical methods while the latter are dependent on the normality assumption for their validity.

The first estimation method is an empirical technique and is referred to as the resubstitution method. C. A. B. Smith (1947) suggested that the sample used to compute the discriminant function could be reused to estimate the error or probability of misclassification. This method consists of classifying each member of the sample of size n_1 from population $1(G_1)$ and the sample of size n_2 from population $2(G_2)$ according to the discriminant function determined by the sample estimators. If m_1 is the number of observations from G_1 which are classified into G_2, and m_2 is the number of observations from G_2 which are classified in G_1, then $\hat{P}r(2|1) = m_1/n_1$ and $\hat{P}r(1|2) = m_2/n_2$.

Lachenbruch and Ray (1968) found in their study that this technique was quite misleading because of the badly biased estimate of the probability of misclassification. If the sample used to compute the discriminant function is not large, this method gives too optimistic an estimate of the probabilities of misclassification.

The second widely used technique is an empirical method called the holdout method. This method consists of dividing the sample of n_1 observations from G_1 into two subsamples. Similarly, the sample of n_2 observations from G_2 are divided into two subsamples. The members of the first subsample of each pair are used to calculate the discriminant function and the classification procedure, while the members of the second subsample of each pair are classified according to that procedure. The proportions of misclassified observations are the desired estimates for the probabilities of misclassification. This method has the advantage of producing unbiased estimators, but these estimators have larger variances than those obtained by the previous method.

There are several drawbacks to this method. First, in many applications large samples are not available. This is especially true in operational testing uses when the data can be expensive in terms of time and money and difficult to obtain. Second, this method is quite uneconomical with data. A larger sample than is necessary to obtain a good discriminant function must be selected to obtain estimates of performance. Third, there are problems connected with the size of the holdout sample. If it is large, a good estimate of the performance of the discriminant function will be obtained, but that function is likely to be poor. If the holdout sample is small, the discriminant function will be better, but the estimate of its performance will be highly variable.

In order to reduce the effect of this last shortcoming, a modification to this technique has been suggested. After the estimates of the probabilities of misclassification have been obtained, recompute the discriminant function using the entire sample. The estimates of performance will remain the same, but the discriminant function will be better. The problem with this modification is that the discriminant function evaluated is not the one used. There may be considerable difference in the performance of the two.

Another method is to replace the population parameters by their estimates from the samples; that is, replace $\underset{\sim}{\mu}_i$ by $\overline{\underset{\sim}{X}}_i$ and $\underset{\sim}{\Sigma}$ by $\underset{\sim}{S}^*$. For normally distributed variables with known parameters, the probability of misclassification is

$$\Phi\left(-\frac{\Delta}{2}\right)$$

where $\Delta^2 = (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)' \underset{\sim}{\Sigma}^{-1}(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$

Using estimated parameters, the estimated probability of misclassification becomes

$$\Phi\left(-\frac{D}{2}\right) ,$$

where $D^2 = (\overline{\underset{\sim}{X}}_1 - \overline{\underset{\sim}{X}}_2)' \underset{\sim}{S}_*^{-1}(\overline{\underset{\sim}{X}}_1 - \overline{\underset{\sim}{X}}_2)$ is the Mahalanobis sample distance. If the degrees of freedom are large, this is a fairly accurate estimate of P_i since D^2 is consistent for Δ^2, the population Mahalanobis distance. If the degrees of freedom are not large, this may be badly biased and give much too favorable an estimate of the probability of error.

Finally, a method was suggested by Lachenbruch (1967) that is an empirical method that decreases the seriousness of bias while making use of all the observations. This procedure retains all of the advantages of both the resubstitution method and the holdout method. From the first sample, exclude the first observation and compute the discriminant function on the basis of the remaining observations. Then classify the excluded observation. Do this for each member of the first sample. The proportion of misclassified observations estimates $Pr(2|1)$. A similar process applied to the second sample estimates $Pr(1|2)$ error rates for a discriminant function based on $n_1 - 1$ and n_2 observations. The estimates of the probabilities of misclassification are then computed by summing the number of cases that were misclassified from each group and dividing by the number in each group. On the basis of Monte Carlo studies, Lachenbruch and Mickey have shown that the bias of these estimates is negligible.

The disadvantage of this method is the computational effort involved. This method requires the computation of $n_1 + n_2$ discriminant functions with an equivalent number of matrix inversions. Another disadvantage is that the actual discriminant function using the entire sample is not evaluated.

The method used in this paper to estimate the probabilities of misclassification can be considered as an amalgam of the resubstitution and holdout empirical methods with computer simulation. As in the resubstitution method, first determine the discriminant function using all of the sample observations from both samples. Then by using computer simulation, generate 10,000 observations from each population with the use of the sample parameters. Then as in the holdout method, classify each of the generated observations into one of the two populations. The probabilities of misclassification will be the number misclassified from each generated sample divided by 10,000. Although the technique of generating these observations is integral to this method, it will be discussed in a later part of this chapter.

This empirical computer simulation technique combines the advantages of both the holdout and resubstitution methods. First, the complete sample of observations is used to determine the discriminant function. Therefore, no information is lost by partitioning the sample. Thus it is as economical with data as the resubstitution method and more economical with data than the holdout method. Secondly, this technique produces unbiased estimators of the probabilities of misclassification method. Yet, these unbiased estimators do not have larger variances because the entire sample is used in determining the discriminant function. Thirdly, as opposed to the holdout modification the discriminant function that is evaluated is the discriminant function that is eventually used.

Not unlike the Lachenbruch (1967) method, the computational effort involved in using the simulation method is sizeable. But the use of high-speed computers and the fact that only one discriminant function is determined thus necessitating a maximum of two matrix inversions, negates this disadvantage. The major effort is involved in the generating and classification of the twenty thousand pseudo-observations. This number of observations can be reduced, but as will be seen in the example, the runs on problems with eighteen variables required approximately two minutes total computer time with linear discrimination.

Minimization of Expected Total Cost

As was seen earlier, if the costs of misclassifications are equal, the objective of the discriminant function is to minimize the expected probability of misclassification. When the costs are not equal, the objective of the discriminant function is to minimize the expected total cost of misclassifications.

$$\text{Total cost} = q_1\, C(2|1)\, \hat{P}r\,(2|1) + q_2\, C(1|2)\, \hat{P}r(1|2)$$

Hence a minimization of the probabilities of misclassification will not always minimize the expected total cost of misclassification. A higher probability for one population may be tolerated because of the lower cost of misclassification for that population. Since sample estimators and the empirical computer simulation technique are used in this methodology, new classification rules must be advanced to insure the minimization of the expected total costs of misclassification.

In the case of the linear discriminant function, the classification rule is to classify a generated observation y_{it} such that

$$y_{it} \geq y_c^* \quad \text{classify into Pop K}$$

$$y_{it} < y_c^* \quad \text{classify into Pop j} \;,$$

where $K \neq j$ and y_{it} is the tth observation generated from the ith population. By moving y_c in one direction or the other, the probabilities of misclassification will be changed.

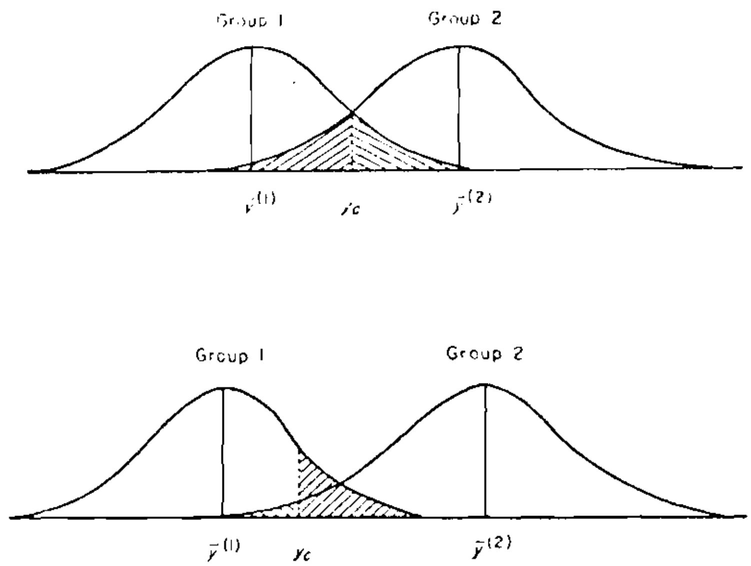

Figure 2-4. Probability of Misclassification with Different y_c [38, p. 247]

Hence in the empirical computer simulation technique, the optimal cutoff point y* will be located at the y_c which yields the lowest expected total cost of misclassification. This optimal cutoff point y* is found by searching the line between the univariate discriminant indices $\overline{Y}_1$ and $\overline{Y}_2$. In the simulation this is accomplished by initially taking y_c to be the midpoint between $\overline{Y}_1$ and $\overline{Y}_2$ and classifying all of the generated observations. From this the probabilities of misclassification are determined, and the expected total cost of misclassification is calculated. At this time y_c is moved to the right a delta increment, and once again all of the generated observations are classified, resulting in the new probabilities of misclassification. The new expected total cost of misclassification is calculated and compared to the previous minimum cost. If the new cost is less, y_c is accelerated in this direction. If the new cost is greater y_c is decelerated in the opposite direction. This linear search of the nonlinear expected cost function will terminate when the change in costs is less than some small distance from the previous minimum. This assures a minimum expected total cost of misclassification plus or minus some small increment and the optimal cutoff point y* which determines it. This is caused by the sum of the probabilities of misclassification being unimodal. The multiplication of these probabilities by constants will translate the curve, but it will remain unimodal.

In the case of the nonlinear discriminant function with equal prior probabilities and costs of misclassification, the classification rule is:

$$\text{Classify into J if } \frac{\hat{P}r(H_j|\underset{\sim}{X}_i)}{\hat{P}r(H_k|\underset{\sim}{X}_i)} \geq 1 \qquad \begin{array}{l} k = 1,2 \\ k \neq j \end{array}$$

As in the linear discriminant function, the objective is to minimize the expected total costs of misclassification. If the priors are not equal or the costs are unequal, the minimization of the probability of misclassification may not be optimal. In the empirical computer simulation technique the classification rule used is:

$$\text{Classify into J if } \frac{\hat{P}r(H_j|\underset{\sim}{X}_i)}{\hat{P}r(H_k|\underset{\sim}{X}_i)} \geq \text{ZETA} \qquad \begin{array}{l} k = 1,2 \\ k \neq j \end{array}$$

If ZETA is equal to one, this procedure is the same as the earlier procedure. Thus the probability that an observation comes from population j has to be greater than the probability that it comes from population K multiplied by ZETA. In much the same manner as in the linear discrimination, the value of ZETA can be changed thereby affecting the classification of the generated observations. When the minimum expected cost of misclassification is obtained, the appropriate ZETA* will be known, and the classification rule will be optimum.

Generation of Sample Observations

The empirical computer simulation technique derives most of its advantages from its ability to generate multivariate normal observations. These observations must exhibit the unbiased estimates of the population parameters that are determined from the sample data. The technique used to generate these observations is the technique developed by Ernest M. Scheuer and David S. Stoller [42]. To generate an observation vector, $\underline{X} = (x_1, x_2, \ldots, x_n)$ with mean vector zero and covariance matrix, Scheuer and Stoller make use of a fundamental theorem of multivariate statistical analysis which states that, if $\underset{\sim}{Z}' = (z_1, \ldots, z_p)$ is multivariate normal with zero mean vector and covariance matrix

$$\underset{\sim}{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ 0 & \cdot & \cdot & \cdot & 1 \end{bmatrix}$$

then $\underset{\sim}{X}$ with mean vector μ and covariance matrix $\underset{\sim}{\Sigma}$ can be represented as $\underset{\sim}{X} = \underset{\sim}{C}\underset{\sim}{Z}$ where C is a lower triangular matrix satisfying $\underset{\sim}{\Sigma} = \underset{\sim}{C}\,\underset{\sim}{C}'$ [4]. Then $\underset{\sim}{X}$ is distributed $N(\underset{\sim}{\phi}, \underset{\sim}{C}\underset{\sim}{C}')$ or $N(\underset{\sim}{\phi}, \underset{\sim}{\Sigma})$. The assumption that the mean vector of the generated observations is the zero vector entails no loss of generality, for if a vector $\underset{\sim}{X}$ with mean vector zero and covariance matrix $\underset{\sim}{\Sigma}$, the vector $\underset{\sim}{X} + \underset{\sim}{U}$ has the same covariance $\underset{\sim}{\Sigma}$ and mean vector $\underset{\sim}{U}$. Hence the generation of the observation takes three steps: (1) compute lower triangular matrix $\underset{\sim}{C}$, (2) generate p independent normal variates (Z), (3) determine the observation vector $\underline{X}$.

In order to compute the unique and lower triangular matrix $\underset{\sim}{C}$, the elements of $\underset{\sim}{C}$ are determined recursively as follows:

$$c_{i1} = \sigma_{i1}/\sqrt{\sigma_{11}} \qquad 1 \leq i \leq n$$

$$c_{ii} = \sqrt{\sigma_{ii} - \sum_{k=1}^{i-1} c_{ik}^2} \qquad 1 < i \leq n$$

$$c_{ij} = [\sigma_{ij} - \sum_{k=1}^{j-1} c_{ik}c_{jk}]/c_{ij} \qquad 1 < j < i \leq n$$

$$c_{ij} = 0 \qquad i < j \leq n$$

This method of computing $\underset{\sim}{C}$ is the "square root" method described by Faddeeva (1959) and is well suited for computation.

The next step is to generate the $\underset{\sim}{Z}$ vector which is distributed N(0,I). This vector consists of p independent standard normal variables. Tables of such random variables exist and there are methods of machine generation of such variables. The technique used in this simulation is attributed to Box and Mueller (1958). This method generates a pair of random deviates from the same normal distribution starting from a pair of random numbers.

Let U_j, U_{j+1} be independent random variables from the same rectangular density on the interval (0,1). Generate variates from N(0,1) by:

$$Z_j = (-2 \log U_j)^{\frac{1}{2}} \cos(2\pi\, U_{j+1})$$

$$Z_{j+1} = (-2 \log U_j)^{\frac{1}{2}} \sin(2\pi\, U_{j+1})$$

Hence by repetition of the above procedure the $\underset{\sim}{Z}$ vector is obtained.

Finally, the observation vector must be generated. Once the C_{ij}'s and Z_i's have been determined, the entries of the observation are:

$$X_i = \sum_{j=1}^{k} c_{ij} Z_j + \mu_i$$

or

$$\underset{\sim}{X} = \underset{\sim}{C}\, \underset{\sim}{Z} + \underset{\sim}{\mu}$$

Data Analysis

All of the multivariate techniques presented in this methodology are based on the assumption of a multivariate normal distribution for each population that has been sampled. Three classes of distributions that are derivative from a parent multivariate normal distribution are: (1) marginal distributions, (2) conditional distributions, (3) component distributions.

A marginal distribution is the univariate distribution for any single element of a vector variable. If the vector variable is multivariate normal in distribution, then every one of its marginal distributions is normal. However, even if all the marginals are normal, it is not necessarily true that the vector variable is multivariate normal.

A conditional distribution is the predicted distribution for a particular marginal element given the known distribution of the remainder of the vector variable. An important theorem is that if a vector variable is multivariate normal in distribution, then every conditional distribution defined on it is normal.

A component distribution is the distribution of any arbitrary linear function of a vector variable. If a vector variable is multivariate normal in distribution, then every component defined on it is normal. Since conditional distributions are special components of a vector variable, the normality of all conditional distributions is a special application of this more general theorem. Likewise, since each marginal of a vector variable is simply that special component defined by setting a unit weight for the assigned element and zero weights for all other elements, the normality of all marginals is a special case of this general theorem. This general theorem is very important because it is reversible. Thus any vector variable for which every possible linear component is normal is a vector variable that is multivariate normal in distribution.

The problem is how to inspect all of the component distributions for normality. Cooley and Lohnes suggest that the only test needed is on the marginals [14]. The marginal distributions should be examined and if necessary be transformed to fit normality. Even though the marginals do not themselves guarantee a multivariate normal distribution, there is no useful test of multivariate normality. In the methodology of this thesis, the requirement for marginal normality is the only check for multivariate normality. If all of the marginals are distributed normally, the vector variable is assumed to have a multivariate normal distribution.

The test used is the Kolmogorov-Smirnov (K-S) test. The K-S test allows the evaluation of the hypothesis that a sample of data was drawn from a specified continuous distribution. The test is nonparametric and exact for all sample sizes in contrast to the asymptotic nature of the chi-square test. Since the distribution under the alternative hypothesis is usually not known, it is not possible to determine the power of the test exactly. Massey (1951) and Kac (1955) compared the power of the K-S test with the power of the chi-square. The K-S test was found to be more powerful in the case of testing for normality with U and σ^2 estimated by $\overline{X}$ and S^2. Let $X_1,\ldots,X_n$ be n observations which are ordered from smallest to largest with sample cumulative distribution function

$$\hat{F}(x) = \begin{array}{ll} 0 & \left\{\begin{array}{l} x < x_1 \\ x_j \leq x < x_{j+1} \qquad j = 1,\ldots,n-1 \\ x \geq x_n \end{array}\right. \end{array}$$

The test statistic is

$$D(x) = \max_x |\hat{F}(x) - F_o(x)|$$

which can be shown to have a distribution that does not depend on $F_o(x)$.

Critical values D_α for significance level α and sample size n can be found in tables by Lilliefors (1967). The null hypothesis is rejected if $D_\alpha < D(x)$.

In the case when the parameters are estimated from the sample, the D values for the K-S test are not exact. Modified critical values for the test statistic have been given by Lilliefors for testing normal distributions with maximum likelihood estimators for the unknown parameters. Use of the K-S test without adjustment for parameter substitution can seriously undermine the validity of any conclusion from the test.

CHAPTER III

DEMONSTRATION OF THE METHODOLOGY

Introduction

In this chapter the methodology developed in Chapter II will be demonstrated. The example used in this demonstration is the actual Operation Test I of the Squad Automatic Weapon System (SAW). This test consisted of observing three SAW candidate weapons in the hands of user/troops during the conduct of Operational Test I. The test was conducted during the period 6 May 1974 to 5 August 1974 at Fort Benning, Georgia. Due to the classification of some data and of the candidate weapons, some characteristics will not be used in this evaluation and the candidates will be referred to as Candidate A, Candidate B, and Candidate C.

The test itself is divided into four major subdivisions: (1) Training Subtest, (2) Quick Fire Subtest, (3) Day Defense Subtest, (4) Attack Subtest. The training subtests are divided into two phases. The first phase was the selection of test soldiers for the remainder of the tests. Forty-eight infantry soldiers possessing desirable characteristics participated in initial training and qualification with the M60 machine gun. The evaluation of the M60 machine gun training and qualification results led to selection of 24 soldiers to continue the test. The second phase of the training subtest evaluated the type, duration, and adequacy of instructions given by the test control personnel to the test soldiers on each candidate system.

The Quick-Fire Subtest provided accuracy and operational effectiveness data on each candidate weapon system in a daylight quick-fire environment. Test soldiers traversed the quick-fire course three times, firing a different weapon system each time. All firing was in the automatic mode from the shoulder using the pointing technique.

The Day Defense Subtest provided accuracy and operational effectiveness data on each candidate weapon system in the daylight defensive role. The test soldiers fired with each weapon system at timed targets of differing ranges. All test soldiers fired this subtest from the foxhole supported position with the hasty mount. During the overall sequence of engagement of the targets, the firer would expend a magazine of ammunition necessitating a timed magazine change. After the firing sequence, each firer was required to change the weapon's barrel in a timed exercise while in the foxhole supported position.

The Attack Subtest provided accuracy and operational effectiveness data on each candidate weapon system in the daylight offensive role. Each soldier traversed the course twice with each weapon system (once with a 100-round magazine and once with a 200-round magazine). Each firer traversed the first half of the course in the shoulder fire position and the second half in the hip or sling-supported underarm assault fire position. After completion of the course, the test soldier was timed in a magazine change and a barrel change while in the standing position.

The objective of the Operational Test is to evaluate the operational issues for each candidate weapon and to enable OTEA to make a statement as to preference with respect to these issues. One operational issue is the operational effectiveness of each weapon. The methodology presented in Chapter II will be used to assess the risk in making a preference statement relevant to operational effectiveness for a candidate over another candidate.

The test results contained data on 19 variables that were used in evaluating operational effectiveness (Table 3-1). Each of the 24 test soldiers received a score on each variable for each candidate weapon. Thus the score on the 19 variables for a soldier on each candidate can be combined into an observation vector for that candidate weapon system. Hence, each candidate weapon system is considered a population and the score vector from each soldier on the weapon can be considered a sample observation from the population. Therefore, the problem is initially structured such that there are three populations with 24 observations on each population and each observation consists of 19 variables.

Data Analysis

The first part of the empirical computer simulation technique is the validation of the assumptions of multivariate normality. As discussed in the data analysis portion of Chapter II, the marginals of the weapons distribution were evaluated using the Kolmogorov-Smirnov Test and Lilliefors critical values for the K-S normality test. The results, given by Table 3-2, show the marginal distributions on 18 variables is normal when the data is in original form or transformed by the square root or natural logarithm. The non-normal marginal is associated with X_{19}, the percent of targets engaged.

Variable X_{19} is a percentage, and it is well known that percentage numbers between zero and one can often be transformed to

Table 3-1. Operational Effectiveness Variables--SAW

Variable	Description
	Training Subtest
X(1)	Basic Firing Qualification Scores
X(2)	Transition Firing Scores
X(3)	Time for Disassembly (Sec.)
X(4)	Time for Assembly
	Quick Fire Subtest
X(5)	Avg. Time to First Round (Sec.) - 20 meter target
X(6)	" " " " " " - 40 meter target
X(7)	" " " " " " - 60 meter target
X(8)	" " " " " " - 80 meter target
X(9)	" " " " " " - moving target
	Day Defense Subtest
X(10)	Avg. Time to First Hit (Sec.)
X(11)	Time to Change Magazine (Sec.)
	Attack Subtest
X(12)	Time to First Round - Sling Position - 100 round mag.
X(13)	" " " " " " - 200 round mag.
X(14)	" " " " " " - 100 round mag.
X(15)	" " " " " " - 200 round mag.
X(16)	Magazine Change Time 200 round/100 round
X(17)	" " " 100 round/200 round
X(18)	Barrel Change Time (Sec.)
X(19)	Percent Targets Engaged - Day Defense

Table 3-2. K-S Test

Variables	Candidates	Transformations: None	Transformations: $\sqrt{}$	Transformations: log	Transform Used
X_1	"A"	.1165	.1042	.0973	none
	"B"	.0846	.1043	.1228	
	"C"	.1365	.1688	.2216	
X_2	"A"	.2448	.1621	.3328	$\sqrt{}$
	"B"	.1048	.1438	.3202	
	"C"	.1293	.1681	.2935	
X_3	"A"	.1481	.1321	.1361	none
	"B"	.1167	.0927	.0968	
	"C"	.1664	.1520	.1356	
X_4	"A"	.1765	.1432	.1044	log
	"B"	.1953	.1567	.1138	
	"C"	.2244	.1894	.1548	
X_5	"A"	.0917	.0824	.0734	none
	"B"	.0887	.1036	.1191	
	"C"	.0730	.0892	.1048	
X_6	"A"	.2495	.2097	.1743	log
	"B"	.1242	.1123	.1040	
	"C"	.2240	.2000	.1772	
X_7	"A"	.2077	.1884	.1800	log
	"B"	.1784	.1614	.1472	
	"C"	.2397	.2062	.1795	
X_8	"A"	.1465	.2419	.3492	none
	"B"	.1842	.1620	.1551	
	"C"	.1529	.1370	.1214	
X_9	"A"	.1406	.1752	.3592	none
	"B"	.1378	.1853	.3287	
	"C"	.1251	.1905	.3485	
X_{10}	"A"	.0661	.0773	.1244	none
	"B"	.1229	.1573	.1909	
	"C"	.1197	.1184	.1373	
X_{11}	"A"	.1934	.1582	.1909	$\sqrt{}$
	"B"	.1868	.1510	.0994	
	"C"	.1402	.1132	.0994	

Table 3-2. Continued

Variables	Candidates	Transformations: None	Transformations: $\sqrt{\ }$	Transformations: log	Transform Used
X_{12}	"A"	.2242	.1743	.1227	log
	"B"	.1576	.1200	.1053	
	"C"	.2096	.1771	.1364	
X_{13}	"A"	.3067	.2351	.1517	log
	"B"	.2491	.1721	.1546	
	"C"	.2096	.1771	.1361	
X_{14}	"A"	.2775	.1874	.0982	log
	"B"	.0971	.0855	.1839	
	"C"	.2148	.1610	.1295	
X_{15}	"A"	.2574	.1957	.1256	log
	"B"	.1928	.1609	.1189	
	"C"	.1812	.1511	.1152	
X_{16}	"A"	.2682	.2268	.1747	log
	"B"	.1725	.1442	.1477	
	"C"	.2614	.2126	.1555	
X_{17}	"A"	.1718	.1168	.1025	log
	"B"	.1604	.1280	.1218	
	"C"	.3287	.2397	.1588	
X_{18}	"A"	.2887	.2039	.1291	log
	"B"	.1113	.0952	.0800	
	"C"	.1948	.1449	.0958	
X_{19}	"A"	.2354	.2343	.2477	.2749*
	"B"	.2883	.2851	.2816	.3066
	"C"	.1517	.1550	.1556	.2020

*Arc sin

normality through the Arcsin transformation. Thus this variable was transformed under the Arcsin and tested. Normality was again rejected. An analysis of the data showed that with one candidate weapon, 13 out of 24 test soldiers engaged 100 percent of the targets while with another candidate 12 soldiers engaged 100 percent of the targets. Thus, not only is the data truncated, but over half of the observations are on the end point. This led to exclusion of this variable from the observation vector.

Therefore, the observation space was collapsed to 18 variables. The assumption of multivariate normality was justified because of the normality of the marginal distribution of each of the remaining variables.

In order to determine the risk involved with a preference statement, each pair of candidates was assembled and tested. In the linear discrimination procedure, the assumption of homoscedasity (equality of covariance matrices) must be used. The 24 observations on each candidate were used to determine the sample means and the pooled-within sample covariance matrix for each two-candidate population. Thus we have

$$\overline{X}_{ij} = \frac{1}{24} \sum_{k=1}^{24} x_{jk} \qquad j = 1,\ldots,18$$

$$\overline{\underset{\sim}{X}}_{i} = (\overline{x}_{j1}, \overline{x}_{i2}, \ldots, \overline{x}_{i18}),$$

and

$$S_* = \{s_{i_j}\} \qquad s_{i_j} = \frac{1}{24+24-2} \sum_{m=1}^{2} \sum_{k=1}^{24} (x_{m_{k_i}} - \bar{x}_{m_i})(x_{m_{k_j}} - \bar{x}_{m_j})$$

$$i = 1,\ldots,18$$
$$j = 1,2,\ldots,18$$

There are many computer programs available to perform the computations for the estimated sample means and pooled-within sample covariances. The observations were entered into the Biomedical Computer Program BMD07M. The means of each sample and the pooled correlation matrices are shown in Appendix I.

The assumption of equal dispersion could have been made with reference to all three populations at one time. This would have necessitated the calculation of only one pooled-within sample covariance matrix. This was not done because each candidate is being compared with one other candidate. The single pooled covariance matrix would dilute the quality of the estimation which would lead to a loss in discriminating power of the empirical computer simulation technique.

Results of Empirical Computer Simulation Technique

The next step in the methodology is to find the optimal linear discriminant function, generate sample observations, classify the generated observations, and compute the probability of misclassification. In Appendix II, there is an interactive Fortran Computer Program called MISSCLASS. This program, with the means of the two samples and the pooled-within population covariance matrix:

1. Calculates the weighting coefficients for the linear discriminant function.

2. Generates ten thousand observations from each population.
3. Classifies all generated observations.
4. Optimizes the cutting point by minimizing the expected total cost of misclassification.
5. Determines the optimal probabilities of misclassification.

In this problem the prior probabilities of observations coming from specific populations were considered equal. The prior probabilities in the context of this specific example are meaningless.

The costs of misclassification were also assumed to be equal. The costs in this problem can be viewed as the driving force in the optimization phase. A value of 1000 was entered into the MISSCLASS Program for each cost. A cost of this magnitude and a search interval of one, causes the program to search until it is within one dollar of the minimum expected cost. Therefore,

$$q_1 = q_2 = .5$$

$$C(1|2) = C(2|1) = 1000$$

The results of the runs of all three combinations of candidates are shown in Table 3-3. As can be determined by examining the probabilities of misclassification, all three candidate weapons show very different operational effectiveness characteristics. Candidates A and C seem to resemble each other the most, but a probability of misclassification of approximately .02 cannot be interpreted as showing great similarity. The seeds used in the random number generator can cause small perturbations in the results, although the size of the generated sample should smooth this out. Another source of error is

Table 3-3. Results of Empirical Computer Simulation Technique

"A" vs "B"	
Linear Discrimination	
Probability of misclassifying an observation from "A" =	.0098
" " " " " " "B" =	.0068
Computer time = 119826 MLSEC	
Nonlinear Discrimination (Equal Dispersion)	
Probability of misclassifying an observation from "A" =	.0096
" " " " " " "B" =	.0079
Computer time = 257560 MLSEC	
Nonlinear Discrimination (Nonequal Dispersion)	
Probability of misclassifying an observation from "A" =	.0011
" " " " " " "B" =	.0003
"A" vs "C"	
Linear Discrimination	
Probability of misclassifying an observation from "A" =	.0233
" " " " " " "C" =	.0187
Computer time = 122674 MLSEC	
Nonlinear Discrimination (Equal Dispersion)	
Probability of misclassifying an observation from "A" =	.0232
" " " " " " "C" =	.0193
Computer time = 258831 MLSEC	
Nonlinear Discrimination (Nonequal Dispersion)	
Probability of misclassifying an observation from "A" =	.0005
" " " " " " "C" =	.0004
"B" vs "C"	
Linear Discrimination	
Probability of misclassifying an observation from "B" =	.0009
" " " " " " "C" =	.0003
Computer time = 125792 MLSEC	
Nonlinear Discrimination (Equal Dispersion)	
Probability of misclassifying an observation from "B" =	.0009
" " " " " " "C" =	.0003
Computer time = 261157 MLSEC	
Nonlinear Discrimination (Nonequal Dispersion)	
Probability of misclassifying an observation from "B" =	.0000
" " " " " " "C" =	.0006

in the matrix inversion routine which is subject to roundoff errors. In the test of Candidate "A" against Candidate "C" it is noteworthy to remember that 233 generated observations out of 10,000 from "A" were classified into "C". This is opposed to the test between "B" and "C" where only three generated observations out of 10,000 from "C" were classified into "B".

The interpretation of the overall results of the linear discrimination should be that a preference statement regarding "B" and "C" has very little risk of being reversed while a preference statement regarding "A" and "C" has a great deal more risk of being reversed though this risk is still not very large.

Another important area of information is the weighting coefficients for the linear discriminant function. Table 3-4 shows the standardized discriminant weights for the three comparisons and Table 3-5 shows these coefficients ranked in order of magnitude.

In the "A" versus "B" competition, variables X(16) and X(17) have relatively great standardized discriminant weights. These variables are associated with the magazine change times in the Attack Subtest--X(16) is the change time from a 200-round magazine to a 100-round magazine, while X(17) is the time in the opposite sequence. These standardized discriminate weights mean that there is a lot of difference between the two weapons with respect to a soldier's ability to change magazines in the upright attack position.

In comparing the same weapons system, there are six standardized weights that have an absolute value less than one. This is interpreted as meaning that there is relatively little difference between these

Table 3-4. Standardized Weighting Coefficients for Discrimination

Coefficient	Test		
	"A" vs "B"	"A" vs "C"	"B" vs "C"
A(1)	-2.36	2.40	2.83
A(2)	2.23	1.60	.58
A(3)	.29	2.30	- .08
A(4)	2.57	1.66	- .81
A(5)	-2.05	-2.28	-2.85
A(6)	.64	- .02	.60
A(7)	1.80	- .98	-2.13
A(8)	-2.09	1.53	-1.49
A(9)	1.35	- .21	-2.84
A(10)	.67	2.42	1.47
A(11)	.21	4.24	5.66
A(12)	-1.17	-1.21	1.41
A(13)	.80	.92	-1.43
A(14)	-1.79	.48	-1.47
A(15)	.25	1.94	2.22
A(16)	-2.93	1.90	7.53
A(17)	-3.64	- .20	.87
A(18)	-1.26	2.58	1.77

Table 3-5. Discrimination Coefficients Ranked in Order of Magnitude

"A" vs "B"	"A" vs "C"	"B" vs "C"
A(17)	A(11)	A(16)
A(16)	A(18)	A(11)
A(4)	A(10)	A(5)
A(1)	A(1)	A(9)
A(2)	A(3)	A(1)
A(8)	A(5)	A(15)
A(5)	A(15)	A(7)
A(7)	A(16)	A(18)
A(14)	A(4)	A(8)
A(9)	A(2)	A(10)
A(18)	A(8)	A(14)
A(12)	A(12)	A(13)
A(13)	A(7)	A(12)
A(10)	A(13)	A(17)
A(6)	A(14)	A(4)
A(3)	A(9)	A(6)
A(15)	A(17)	A(2)
A(11)	A(6)	A(3)

weapons on the six operational effectiveness characteristics. An interesting point is revealed during examination of variables X(12), X(13), X(14), and X(15). Variables X(13) and X(15) have absolute standardized weights less than 1.0, whereas X(12) and X(14) have absolute standardized weights greater than one. This may signify that the two weapons are operationally more similar with the 200-round magazine, X(13) and X(15), than they are with the 100-round magazine, X(12) and X(14), in the supported firing position during the Day Defense.

In the comparison of candidate "A" and "C", the standardized weight given to variable X(11), time to change magazine in Day Defense, is much greater in absolute value than any other weight. Thus the operational characteristic of magazine change time in this position seems to be very different between the two weapons.

The signs of the standardized discriminant coefficients also can be used to make inferences about the competing systems. As an example, a negative sign will infer that the mean of the first population is less than the mean of the second population on this variable. The benefit of this sign inference to a candidate is not standard, because on some variables low means are desired while on others high means are desired.

Three variables in the Quickfire subtest show little discriminating ability. Variables X(6), X(7), and X(9) have absolute standardized weights of less than 1.0. This tends to infer that in the time to first round characteristic of a surprise target situation, both candidates are quite similar.

In the comparison of candidate "B" and "C", two variables are weighted very heavy in the discrimination, X(11) and X(16). Both of these variables are concerned with magazine change times. Therefore, the inference may be made that there is a great difference between the two weapons in the soldier's ability to change magazines.

In the Quickfire Subtest three of five measured variables (X(5), X(7), X(9)) have magnitudes greater than 2.0 with only one variable, X(6), being less than 1.0. The inference to be made is that these two weapons differ substantially in the ability of the firer to discharge the first round after target recognition in a surprise type environment. The only positively signed variable in this group is X(6), which is very close to 0.0 in weight. This same inference cannot be made regarding the time to first round characteristic in the Attack Subtest. Although the magnitudes of the weights are all greater than 1.0, the sign changes do not allow for general inferences.

The same comparisons were carried out using the nonlinear discrimination programs F CLASS. Under the assumption of common dispersion matrices, this type of discrimination is equivalent to linear discrimination. As can be seen in Table 3-3, the probabilities of misclassification are almost identical for the two types of discrimination. What little difference there is can be attributed to roundoff errors in the computer. The interesting fact is that the nonlinear discrimination process takes significantly more time to accomplish than the linear discrimination. This is caused by the number of matrix multiplications needed to classify one observation in nonlinear discrimination. In order to determine the density value of an observation with respect to a single population entails two matrix multiplications. Thus four matrix multiplications must be accomplished before an observation can be classified. Since 20,000 observations must be classified, 80,000 matrix multiplications must be done. The result is that linear discrimination is much more efficient and should be used when equal dispersion matrices can be assumed.

The Squad Automatic Weapon data was also used in a nonlinear discrimination with the assumption of equal dispersion relaxed. The results are shown in Table 3-3. Even though the probabilities decreased in every instance, any statement made regarding inequality of dispersion would be subject to much misgiving. This is because the values of the probabilities were very small before the assumption was relaxed. In examining the results though, the decrease in the "A" versus "C" comparison was much more substantial than the rest. With equal dispersion assumed, candidates "A" and "C" were most similar. With the assumption relaxed "A" and "C" are no more similar than the other candidates. Thus a case may be made for not assuming equal dispersion for "A" and "C".

The result of the example problem is that the Squad Automatic Weapon candidates are all different in operational characteristics. The risk in making a preference statement with respect to operational characteristics of one candidate over the other would be small. Variables X(6) was never assigned a heavy weighting coefficient and could thus be assumed to be equivalent in all candidates. On the other hand, X(11) was assigned a heavy weight in two tests, but not in the third. This infers that the magazine change times in the day defense were different in the "A" and "C" test and the "B" and "C" test but not in the "A" and "B" test. Therefore, the magazine change time must be very similar in candidates "A" and "B" but different in candidate "C".

One other interesting fact appears in the results. The Biomedical Computer Program BMD07M used to determine the pooled-within population covariance matrix also discriminates and reclassifies all of the original observations, a resubstitution method. The probabilities of misclassification from BMD07M are shown in Table 3-6. As can be seen the resubstitution method was overly optimistic on four of the six probabilities, equal on one probability, and was greater than the empirical computer simulation technique on only one probability of misclassification.

Table 3-6. Results from BMD07M Classifications

"A" vs "B"	
Probability of misclassifying an observation from "A"	= .000
" " " " " " "B"	= .000
"A" vs "C"	
Probability of misclassifying an observation from "A"	= .0417
" " " " " " "C"	= .0000
"B" vs "C"	
Probability of misclassifying an observation from "B"	= .0000
" " " " " " "C"	= .0000

Comparison of Methods

At the present time, the methodology used to evaluate the SAW test data consisted of a series of Duncan's Multiple Range tests for each variable. From these tests, the candidates were compared variable by variable with significant differences noted.

In many cases the proposed multivariate methodology gave the same results as achieved by the multiple range tests. According to the range test approach, soldier transition firing scores (X_2) with candidates "B" and "C" were significantly higher than the scores with "A". This finding is reinforced by the linear discriminant approach through the analysis of the standardized weighting coefficients. In the "A" versus "B" and the "A" versus "C" discriminations, the weighting coefficients (β_2) are greater in magnitude than 1.6; whereas, in the "B" versus "C" discrimination β_2 is equal to .58. Therefore the conclusion reached is that "B" and "C" differ much more from "A" than they do from each other.

The time to change magazines in the day defense (X_{11}) is another variable that was deemed significantly different for the candidates by the range tests. The mean times with candidates "B" and "A" were declared significantly less than with "C". Again the standardized weighting coefficients confirm this finding. The weights for X_{11} in "A" versus "C" and "B" versus "C" are greater than 4.24 while β_{11} for "A" versus "B" is only .21. Hence the finding under the proposed methodology is that candidates "A" and "B" are very similar with respect to X_{11} but show great difference to "C" on this variable.

Another variable of agreement is X_{16}. The present methodology found that the magazine change with "B" took significantly less time than with "A" and "A" significantly less than "C". The interpretation of the weighting coefficients concurs with this conclusion. All three candidates are shown to be very different by the magnitude of β_{16} in all three linear discriminant functions.

One area where the two methodologies differ is the significance of the time to first hit (X_{10}) in the Day Defense Subtest. The multiple range tests conclude that there was no significant difference between the three candidates. An examination of the β_{10} weights revealed that the weights were much greater in the "A" versus "C" and "B" versus "C" tests than the β_{10} in the "A" versus "B" test. Hence the time to first hit in the Day Defense is much more important in discriminating between "A" or "B" and "C" than it is in discriminating between "A" and "B".

Another area where the multivariate methodology inference seems to disagree with the multiple range tests methodology is the time to first round in the Quickfire Subtest. The range test found no significant difference between the three candidates on variables X_5, X_6, X_7, X_8, and X_9. The proposed procedure never assigned a heavy weight to X_6 but did assign moderate values to β_5, β_7, β_8, and β_9. Especially important to discrimination are variables X_5 and X_8 which are assigned standardized weights greater than 2.05 and 1.49 respectively. On the time to fire first round on moving target (X_9), "A" and "C" show great similarity while neither shows much similarity with candidate "B".

The variable X_1 also deserves an examination due to its uniformly high absolute weights in all comparisons. The Duncan's Multiple range test did not declare a significant difference between the candidates on this variable. This is somewhat contrary to the discriminant weights, which identifies this variable as one of the major areas of differences between the candidates.

Thus, the proposed multivariate methodology does identify the same areas of difference that the multiple range test procedure. The proposed methodology also identifies areas that are important in identifying differences even though not significant under the Duncan's Multiple Range tests. Therefore, a major advantage of the proposed methodology is that it provides at least the same information with considerably less calculations.

Another advantage of the new procedure is that there is a final aggregate measure with which to compare the candidates, the probability of misclassification. The present system does not give a comparison of an entire system with another. Thus the significant differences must be combined in some subjective manner to compare systems. The proposed methodology gives the needed quantitative aggregate measure on which a similarity statement can be made.

CHAPTER IV

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

The definition of "risk" is unclear and changes as the type of problem referenced changes. In the class of problems where candidate prototype pieces of equipment are tested resulting in preference statements, risk can be defined as the probability of making an incorrect preference statement. In order to quantitatively determine the value of the risk inherent in the preference statement, the probability of misclassification can be used as an estimator.

This procedure and definition of risk can also be used with problems dealing with improved equipment opposed to base-line equipment. The variables needed to evaluate the operational issue can be measured on both the base-line and improved equipment. The expected cost of misclassification can be estimated, thus estimating the risk involved in accepting or not accepting the new equipment.

The probability of misclassification can be used in other more subtle areas of operational testing. The selection of test soldiers is a very fertile field for this procedure. Multivariate techniques have been used for years in personnel selection and can be applied to the selection of operational test personnel. The probability of misclassification can be used as a measure of risk in selecting the desired personnel for the test.

Another conclusion of this work is that the empirical computer simulation technique is a viable method of estimating the probability of misclassification. This technique obtains the best population parameter estimates given a sample, formulates the best discriminating function, and estimates the probability of misclassification of the discriminant function empirically. The empirical computer simulation technique takes advantage of the high speed of modern computing machines and the art of simulation in order to efficiently estimate the probability of misclassification.

Recommendations

The robustness of the methodology presented in this paper should be examined. The effects of noncompliance to the methodology assumptions of the processes should be evaluated. If needed, other discriminating functions and observation-generating procedures could be introduced to allow the analysis of other than multivariate normal populations.

Another area of future study is the effect of small sample sizes on the empirical computer simulation technique. The sample estimates of the population means and covariance matrices improve as the size of the sample increases, but with small samples, the estimates are subject to question.

Finally the application of this procedure to an enlarged data base composed of successful and nonsuccessful populations should be examined. It is conceivable that values of certain parameters and the relationship between these parameters could be used to define successful and nonsuccessful populations. Whether these parameters, values, and relationships are determined through subjective expert opinion or historical data of like pieces of equipment, the two critical populations could be defined. A new piece of equipment could be evaluated and classified into one of the two populations. Associated with this classification would be a probability of misclassification, thus defining the risk involved with accepting the new equipment.

APPENDIX I

CORRELATION MATRICES FROM SQUAD AUTOMATIC WEAPONS TEST

Candidate "A"

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	.107	1.000																
3	.103	-.129	1.000															
4	.162	-.250	.818	1.000														
5	.165	.132	.219	.119	1.000													
6	.517	-.433	-.070	-.024	-.029	1.000												
7	.039	.349	.001	-.157	.199	.119	1.000											
8	.085	.011	-.038	-.171	.127	-.028	.525	1.000										
9	.120	.139	-.244	-.465	.055	.099	.396	.555	1.000									
10	.022	-.263	-.349	-.403	-.201	.123	-.259	.105	.229	1.000								
11	-.139	.099	-.343	-.372	.173	-.050	-.318	-.326	.185	.334	1.000							
12	.501	.078	.518	.313	-.205	.198	-.033	-.047	-.030	-.062	-.224	1.000						
13	-.257	-.045	.139	.205	.044	-.231	-.172	-.182	-.485	-.228	-.141	-.031	1.000					
14	.204	-.094	-.073	-.104	-.397	.230	-.229	-.346	-.215	.038	-.195	.422	.293	1.000				
15	-.130	-.126	.431	.380	.010	-.042	.292	.142	-.231	-.520	-.545	.134	.135	-.025	1.000			
16	-.171	.279	-.226	-.217	.019	-.141	.450	-.060	.122	-.177	.147	-.134	-.108	.029	.205	1.000		
17	-.569	.007	-.020	-.037	-.262	-.279	-.045	-.204	-.118	-.171	.124	-.271	-.057	-.145	.219	.279	1.000	
18	-.176	.261	-.142	-.184	.040	-.140	.067	-.013	.133	.017	.028	-.276	-.183	-.262	-.199	-.106	.155	1.000

Candidate "B"

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	.198	1.000																
3	-.081	-.452	1.000															
4	.016	-.106	.510	1.000														
5	.038	.225	.071	.159	1.000													
6	.264	-.349	.181	.046	.252	1.000												
7	.081	-.107	-.098	.204	.523	.639	1.000											
8	.281	.275	-.084	-.191	-.075	.083	-.064	1.000										
9	.407	.161	-.099	-.368	.086	.050	-.208	.258	1.000									
10	.582	.416	-.027	.046	-.251	-.126	-.189	.504	.132	1.000								
11	-.275	-.265	-.161	-.501	.001	.107	.117	.064	.431	-.468	1.000							
12	.286	-.064	-.279	-.129	-.431	.096	.111	.029	-.090	.229	-.088	1.000						
13	-.036	-.120	.116	-.109	-.042	.167	-.241	.015	.231	.030	.066	.200	1.000					
14	.054	.305	.107	.148	-.123	-.068	-.338	-.235	-.008	.144	-.405	.025	.498	1.000				
15	.196	.501	-.146	.223	-.012	-.270	-.236	.051	.198	.264	-.278	-.031	.274	.649	1.000			
16	-.125	.006	.292	.090	.165	-.159	-.328	-.089	.264	-.202	-.106	-.275	.217	.129	-.145	1.000		
17	-.312	-.516	-.083	-.005	-.401	-.119	-.006	-.283	-.285	-.190	.104	.325	-.047	-.273	-.133	-.268	1.000	
18	.043	.049	.198	.242	.127	-.262	-.033	-.135	.030	-.105	-.082	-.186	-.119	-.063	-.025	.499	-.218	1.000

Candidate "C"

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	-.187	1.000																
3	-.080	-.275	1.000															
4	-.201	-.162	.565	1.000														
5	.060	.258	.143	-.076	1.000													
6	.221	.185	.038	-.048	.292	1.000												
7	-.155	-.164	.403	.359	.086	.221	1.000											
8	-.364	.029	-.017	.044	.241	.383	.376	1.000										
9	.343	.048	-.202	-.044	-.227	-.051	-.477	-.020	1.000									
10	-.290	-.027	.347	.221	.212	.055	.182	-.006	-.302	1.000								
11	.016	.005	-.361	-.206	.082	-.259	-.048	.032	.094	-.427	1.000							
12	.090	-.224	.319	.327	-.196	.106	.073	-.253	.110	.337	-.269	1.000						
13	-.126	-.070	.089	.108	-.048	.335	-.080	-.009	-.406	-.082	-.266	-.028	1.000					
14	-.089	-.093	-.188	-.192	.199	.138	.015	.120	-.247	-.161	.183	.166	.075	1.000				
15	-.073	.061	-.001	-.165	.000	.459	.042	.013	-.226	.076	-.309	.233	.381	.500	1.000			
16	-.364	.068	.527	.369	.435	.202	.350	.335	-.215	.344	-.296	-.114	.032	.057	-.088	1.000		
17	.263	-.154	-.096	-.179	.213	.281	-.028	.183	-.135	-.286	.141	-.101	.459	.418	.312	.003	1.000	
18	.211	.117	-.043	.022	.199	.084	.052	.001	.217	-.268	-.098	-.017	.067	-.067	.084	-.043	.162	1.000

Candidates "A" and "B" Pooled

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	.152	1.000																
3	-.009	-.347	1.000															
4	.091	-.168	.598	1.000														
5	.104	.184	.119	.140	1.000													
6	.408	-.383	.076	.009	.104	1.000												
7	.050	.160	-.037	-.032	.289	.263	1.000											
8	.181	.164	-.067	-.182	.018	.026	.295	1.000										
9	.237	.146	-.147	-.416	.069	.079	.223	.411	1.000									
10	.319	.163	-.120	-.137	-.229	-.011	-.205	.344	.172	1.000								
11	-.207	-.122	-.215	-.444	.075	.030	-.140	-.100	.299	-.169	1.000							
12	.399	-.001	.007	.085	-.321	.151	.014	-.007	-.056	.105	-.148	1.000						
13	-.139	-.091	.121	.026	-.005	-.025	-.179	-.067	-.127	-.065	-.015	.097	1.000					
14	.108	.175	.063	.055	-.215	.050	-.232	-.268	-.090	.111	-.333	.166	.428	1.000				
15	-.023	.119	.124	.307	.002	-.112	.180	.101	-.100	-.167	-.405	.072	.177	.259	1.000			
16	-.153	.149	.046	-.089	.077	-.146	.260	-.069	.170	-.177	.032	-.186	.036	.073	.112	1.000		
17	-.473	-.193	-.044	-.024	-.300	-.222	-.037	-.222	-.168	-.163	.108	-.054	-.049	-.179	.136	.132	1.000	
18	-.087	.153	.050	.008	.078	-.218	.039	-.068	.094	-.044	-.025	-.235	-.147	-.138	-.146	.096	.042	1.000

Candidates "A" and "C" Pooled

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	-.038	1.000																
3	-.015	-.206	1.000															
4	.017	-.215	.612	1.000														
5	.106	.195	.167	.036	1.000													
6	.368	-.157	-.006	-.031	.136	1.000												
7	-.039	.163	.183	-.017	.144	.154	1.000											
8	-.086	.017	-.024	-.116	.163	.111	.484	1.000										
9	.233	.097	-.211	-.304	-.094	.029	.067	.340	1.000									
10	-.147	-.147	.093	-.154	.030	.090	-.080	.060	-.039	1.000								
11	-.044	.045	-.355	-.270	.116	-.163	-.176	-.148	.130	-.124	1.000							
12	.286	-.064	.383	.309	-.199	.154	.007	-.119	.038	.144	-.247	1.000						
13	-.196	-.053	.101	.180	.006	-.035	-.147	-.138	-.444	-.165	-.180	-.029	1.000					
14	.062	-.093	-.133	-.135	-.097	.190	-.143	-.189	-.229	-.056	.011	.304	.217	1.000				
15	-.102	-.057	.187	.227	.006	.143	.219	.106	-.226	-.267	-.400	.170	.203	.161	1.000			
16	-.273	.178	.239	.004	.246	.024	.400	.084	-.044	.094	-.114	-.124	-.054	.041	.086	1.000		
17	.031	-.093	-.080	-.097	.085	.094	-.028	.017	-.119	-.238	.134	-.140	.194	.206	.229	.077	1.000	
18	.018	.196	-.080	-.108	.122	-.066	.061	-.008	.172	-.124	-.043	-.152	-.094	-.177	-.092	-.075	.145	1.000

Candidates "B" and "C" Pooled

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1.000																	
2	.016	1.000																
3	-.080	-.379	1.000															
4	-.068	-.121	.514	1.000														
5	.051	.234	.107	.064	1.000													
6	.238	-.116	.110	.010	.274	1.000												
7	-.064	-.129	.171	.248	.264	.396	1.000											
8	.023	.206	-.060	-.138	.039	.180	.103	1.000										
9	.368	.105	-.152	-.220	-.094	-.007	-.372	.138	1.000									
10	.145	.256	.130	.098	-.034	-.043	-.002	.349	-.078	1.000								
11	-.103	-.139	-.263	-.362	.047	-.094	.018	.048	.234	-.440	1.000							
12	.176	-.129	.007	.033	-.307	.101	.089	-.065	.019	.274	-.184	1.000						
13	-.062	-.106	.103	-.064	-.041	.207	-.161	.011	.000	.000	-.046	.120	1.000					
14	-.001	.198	.014	.074	-.008	.002	-.181	-.161	-.089	.057	-.176	.068	.419	1.000				
15	.035	.289	-.069	.063	-.005	.137	-.068	.033	-.053	.169	-.295	.112	.283	.545	1.000			
16	-.281	.036	.417	.194	.333	.065	.124	.089	-.049	.093	-.226	-.173	.117	.083	-.107	1.000		
17	.155	-.184	-.078	-.086	.093	.180	-.022	.025	-.143	-.206	.119	-.016	.179	.124	.213	-.031	1.000	
18	.143	.078	.071	.144	.168	-.069	.019	-.076	.140	-.180	-.091	-.094	-.048	-.059	.039	.146	.088	1.000

APPENDIX II

COMPUTER PROGRAMS FOR EMPIRICAL COMPUTER SIMULATION TECHNIQUE

A. MISSCLASS: Linear Discrimination

B. F CLASS: Nonlinear Discrimination

```
C
C     MISSCLASS IS AN EMPIRICAL SIMULATION TECHNIQUE COMPUTER PROGRAM
C                USING LINEAR DISCRIMINATION
C
C
C****************   VARIABLE LIST   ********************************
C
C   ALPHA--ACCELERATION CONSTANT USED IN SEARCH ROUTINE
C   BETA--DECCELERATION CONSTANT USED IN SEARCH ROUTINE
C   BUF--ARRAY USED TO GENERATE OBSERVATIONS (GEN. OBS. WITH MEAN 0.)
C   CANVEC--CANDIDATE VECTOR TO BE CLASSIFIED
C   CLASS(I)--NUMBER OF GENERATED OBSERVATIONS FROM POP. I CLASSIFIED
C             INTO POPULATION I
C   CMAT--ARRAY USED IN OBSERVATION GENERATION--C'C=SIGMA
C   COPY--ARRAY USED IN MATHSTAT TO SOLVE SIMULTANEOUS EQUATIONS
C   D--VECTOR OF MEAN DIFFERENCES BETWEEN THE TWO POPULATIONS
C   DEL--STEP LENGTH FOR SEARCH
C   EPS--INTERVAL IN WHICH THE OPTIMUM COST IS FROM SEARCH OPTIMUM
C   INDEX-- POPULATION IDENTIFIER
C   JD--ARRAY NEEDED FOR MATHPACK CALLS
C   KOUNT--MAX NUMBER OF ITERATIONS TO BE USED IN SEARCH
C   KPOP--POPULATION IN WHICH CANDIDATE VECTOR IS CLASSIFIED
C   LEPSCK--FLAG TO DENOTE END OF SEARCH
C   NK--NUMBER OF OBSERVATIONS GENERATED FROM EACH POPULATION - SET AT 10000
C   NL--NUMBER OF VARIATES COMPOSING AN OBSERVATION
C   NSTART--NUMBER OF STARTS FOR RANDOM NUMBER GENERATOR
C   OLDCST--OPTIMAL EXPECTED COST OF MISCLASSIFICATION
C   PI(I)--PRIOR PROBABILITY THAT AN OBSERVATION COMES FROM POP. I
C   PMSCLS--NON-OPTIMAL PROBABILITY OF MISCLASSIFICATION USED IN SEARCH
C   PRMSCLS(I)--OPTIMAL PROBABILITY OF MISCLASSIFYING AN OBSERVATION FROM POP. I
C   SIGMA--POOLED COVARIANCE MATRIX
C   XBAR--MEAN VECTOR FOR POP.1
C   XVEC--GENERATED OBSERVATION VECTOR
C   YBAR--MEAN VECTOR FOR POP.2
C   Z(I,K)--UNIVARIATE INDEX FOR THE K-TH GENERATED OBSERVATION FROM POP. I
C   ZBARX--MEAN OF THE UNIVARIATE INDEX "Z" FOR POPULATION 1
C   ZBARY--MEAN OF THE UNIVARIATE INDEX "Z" FOR POPULATION 2
C   ZCUT--CLASSIFICATION CUTTOFF POINT
C   ZOPTCT--CLASSIFICATION ZCUT THAT MINIMIZES THE EXPECTED COST OF
C           MISCLASSIFICATION
C   ZSAMBR(I)--MEAN Z FOR OBSERVATIONS GENERATED FROM POP. I
C
C*****************************************************************
C
C
C**************   SUBROUTINE LIST   *************************
C
C   CLASSIFY--CLASSIFIES CANDIDATE VECTORS --CISTY
C   CMAT--COMPUTES CMAT MATRIX
C   MAIN--READS INPUT,PRINTS OUTPUT,CALCULATES PROBABILITY OF MISCLASSIFICATION
C   OPTIMIZE--OPTIMIZES PROB. OF MISCLASSIFICATION SUCH THAT EXPECTED COSTS
C             ARE MINIMIZED --OPTZ
C   XVEC--GENERATES OBSERVATION VECTORS
C   ZCUT--DETERMINES INITIAL ZCUT (MIDWAY BETWEEN THE ZBAR'S) --CUTOFF
C
C*********************************************************************
```

MISSCLASS.MAIN

```
C
C
C*************   FUNCTION LIST   ********************************
C
C   SIMNORM--PRODUCES INDEPENDENT NORMAL VARIATES USING THE BOX AND MUELLER
C           TRANSFORMATION OF UNIFORM(0,1) DEVIATES
C   UNIF--PRODUCES DEVIATES WITH UNIFORM(0,1) DISTRIBUTION BY THE MULTI-
C        PLICATIVE CONGRUENTIAL OVERFLOW METHOD
C
C**********************************************************************
C
C
          DIMENSION D(30), JD(30),V(2),
     1Z(2,10000),ZSAMP(2),ZSAMBR(2),CLASS(2)
          COMMON /ONE/XBAR(30),YBAR(30),T
          COMMON /TWO/SIGMA(30,30)
          COMMON /THREE/CMAT(30,30)
          COMMON /FOUR/XVEC(30),BUF(30),ZVEC(30)
        COMMON /FIVE/ML
        COMMON /SEVEN/COST(2),PI(2),PMSCLS(2)
       COMMON /EIGHT/PRMSCL(2),ALPHA,BETA,EPS,DEL
        COMMON /NINE/ CANVEC( 30),  KPOP ,COPY(30,31)
          DATA/IRES/6HYES   /
      EXTERNAL UNIF,RNORM1,GJR
   49 FORMAT (//,5X,'*****      INPUT      *****')
   50 FORMAT (//,5X,'**  COVARIANCE MATRIX  **')
   51 FORMAT (//,5X,'**  C  MATRIX  **')
   52 FORMAT (//,5X,'**  MEAN VECTOR FOR POPULATION 1  **')
   53 FORMAT (//,5X,'**  MEAN VECTOR FOR POPULATION 2  **')
   57 FORMAT (/,2X,8(1X,F8.4))
   60 FORMAT ( )
   75     FORMAT(/,2X,'ENTER THE PRIOR PROBABILITY THAT AN OBS. COMES FR
     1OM POP. 1')
   77    FORMAT(/,2X,'ENTER PRIOR PROBABILITY THAT AN OBS. COMES FROM PO
     1P. 2')
   80  FORMAT(/,2X,'**  COST OF MISCLASSIFYING AN OBS. FROM POP. 1 =',
     1F10.5)
   81  FORMAT(/,2X,'**  COST OF MISCLASSIFYING AN OBS. FROM POP. 2 =',
     1F10.5)
   82  FORMAT(/,2X,'**  PRIOR PROB. AN OBS. BELONGS TO POP. 1 =',
     1F10.5)
   83  FORMAT(/,2X,'**  PRIOR PROB. AN OBS. BELONGS TO POP. 2 =',
     1F10.5)
C*****     READ INPUT
C**
C**
       WRITE (6,90)
   90    FORMAT(/,2X,'ENTER NUMBER OF STARTS FOR UNIF')
         READ (5,60,END=999) NSTART
       DO 95 KN=1,NSTART
          X = UNIF(A)
   95   CONTINUE
          WRITE(6,97)
   97     FORMAT(/,2X,'ENTER MAX NUMBER OF ITERATIONS FOR SEARCH')
          READ (5,60,END=999) KNT
      WRITE (6,101)
  101  FORMAT (/,2X,'ENTER DIMENSION OF POPULATIONS')
```

```
      READ (5,60,END=999) NL
      WRITE (6,103)
  103  FORMAT (/,2X,'ENTER THE ESTIMATED MEAN VECTOR FOR POPULATION #1')
      READ (5,60,END=999) (XBAR(J),J=1,NL)
      WRITE (6,105)
  105  FORMAT (/,2X,'ENTER THE ESTIMATED MEAN VECTOR FOR POPULATION #2')
      READ(5,60,END=999) (YBAR(J),J=1,NL)
      WRITE (6,107)
  107  FORMAT (/,2X,'ENTER THE ESTIMATED WITHIN POP. COVARIANCE MATRIX')
      READ (5,60,END=999) ((SIGMA(K,J),J=1,NL),K=1,NL)
       WRITE (6,108)
  108     FORMAT(/,2X,'ENTER THE ESTIMATED COST OF MISCLASSIFYING AN
     1OBS. FROM POP. 1')
        READ (5,60,END=999) COST(1)
        WRITE (6,109)
  109     FORMAT(/,2X,'ENTER THE ESTIMATED COST OF MISCLASSIFYING AN
     1OBS. FROM POP. 2')
        READ (5,60,END=999) COST(2)
         WRITE (6,75)
        READ (5,60,END=999) PI(1)
        WRITE (6,77)
        READ (5,60,END=999) PI(2)
C
C
C*****     PRINT INPUT FOR ACCURACY CHECK
C
      WRITE (6,49)
      WRITE (6,52)
      WRITE (6,57) (XBAR(J),J=1,NL)
      WRITE (6,53)
      WRITE (6,57) (YBAR (J),J=1,NL)
      WRITE (6,50)
      DO 110 K = 1,NL
      WRITE (6,57)  (SIGMA(K,J),J=1,NL)
  110     CONTINUE
          WRITE(6,80) COST(1)
          WRITE(6,81) COST(2)
          WRITE(6,82) PI(1)
          WRITE(6,83) PI(2)
C
C*****     CALCULATE DISCRIMINATE FUNCTION COEFFICIENTS (A7S)
C
C
C*****     COMPUTE MEAN DIFFERENCE VECTOR
C
      DO 1100 L=1,NL
      D(L) =YBAR(L) - XBAR(L)
 1100  CONTINUE
C**     SOLVE SIMULTANEOUS EQUATIONS FOR COEF.
      NC = NL + 1
      V(1) = 4.
      DO 1210 K=1,NL
      DO 1200 J=1,NL
      COPY (K,J) = SIGMA (K,J)
 1200  CONTINUE
      COPY (K,NC) = D(K)
 1210  CONTINUE
```

```
          CALL GJR (COPY,31,30,NL,NC,$900,JD,V)
           WRITE (6,1230)
 1230      FORMAT(//,2X,'*** DISCRIMINANT FUNCTION COEF. ',10X,' STANDARD
          1IZED WEIGHTS ***')
           DO 1250 M=1,NL
               STAN = COPY(M,NC)*SQRT(SIGMA(M,M))
               WRITE (6,1240) M,COPY(M,NC),STAN
 1240          FORMAT(/,5X,'**     A(',I2,')=',F10.5,20X,'STANDARDIZED =',F
          110.5)
 1250      CONTINUE
C
C
C*****     COMPUTE MEAN CLASSIFICATION STATISTIC (ZBAR) FOR EACH POPULATION
C
          ZBARX = 0.0
          DO 2100 L=1,NL
          ZBAR = COPY (L,NC) * XBAR(L)
          ZBARX = ZBARX + ZBAR
 2100      CONTINUE
          ZBARY = 0.0
          DO 2200 K=1,NL
          ZBAR = COPY (K,NC) * YBAR(K)
          ZBARY = ZBARY + ZBAR
 2200      CONTINUE
C
C
C*****     DETERMINE THE OPTIMAL CLASSIFICATION Z (ZCUT)
C**
C
          CALL CUTOFF (ZBARX,ZBARY,ZCUT,ZOPTCT,LFPSCK)
              WRITE (6,2900) ZCUT
 2900      FORMAT (//,5X,'ZCUT =',F10.5)
C
C
C*****     GENERATE OBSERVATIONS FROM MULTIVARIATE NORMAL POPULATIONS
C**
C
          CALL CMAT1
C**         GENERATE OBS.
C
          NK = 10000
           DO 4900 I=1,2
 4100      DO 4300  K=1,NK
          CALL XVEC1 (RNORM1,UNIF)
          DO 4200 L=1,NL
              Z(I,K) = Z(I,K) + XVEC(L)*COPY(L,NC)
 4200      CONTINUE
 4300        CONTINUE
C
C
C*****     FIND THE ZBAR FOR THE GENERATED OBSERVATIONS
C**
C
          DO 4600 M=1,NK
          ZSAMP(I) = ZSAMP (I)  + Z(I,M)
 4600      CONTINUE
           ZSAMBR (I) = ZSAMP(I)/NK
```

```
      C
                WRITE (6,4610)I,ZSAMBR(I)
       4610    FORMAT (//,2X,'*****  Z BAR SAMPLE POP. #',I2,'  =',F10.5)
       4900    CONTINUE
      C
      C*****     CLASSIFY EACH GENERATED OBSERVATION INTO A POPULATION
      C**
      C
                ALPHA = 3.
                BETA  = -.5
                DEL = .1
                EPS   = 1.
                KOUNT = 1
       5000  CONTINUE
              IF (ZBARX .LE. ZBARY) GO TO 5050
                 LSIDE = 2
                 KPSIDE = 1
              GO TO 5100
       5050      LSIDE = 1
                 KPSIDE = 2
      5100   DO 5200 M=1,NK
            IF(Z(LSIDE,M) .GT.ZCUT) GO TO 5200
            CLASS(LSIDE) = CLASS(LSIDE) + 1.
      5200   CONTINUE
            PRMSCL(LSIDE) = (NK-CLASS(LSIDE))/NK
       5500  DO 5600 M = 1,NK
                IF (Z(KRSIDE,M) .LT. ZCUT) GO TO 5600
                CLASS(KRSIDE) = CLASS(KRSIDE) + 1.
       5600  CONTINUE
                PRMSCL(KRSIDE) = (NK-CLASS(KRSIDE))/NK
      5900   CONTINUE
      C
      C**   FIND OPTIMAL ZCUT BY EXAMINING THE EXPECTED COST'S OF ERRORS
      C
               IF (LEPSCK .GE. 2) GO TO 7000
                CALL OPTZ(ZCUT,LEPSCK,ZOPTCT)
              CLASS(1) = 0.
              CLASS(2) = 0.
                KOUNT = KOUNT + 1
                IF ( KOUNT .GT. KNT) GO TO 6000
               GO TO 5000
       6000      WRITE (6,6010)
       6010      FORMAT (//,2X,'%%%%%%%%   KOUNT EXCEEDED%%%%%%%')
       7000    CONTINUE
      C
      C**
      C*****     PRINT OUTPUT
      C**
      C
       9000  CONTINUE
            WRITE (6,9100)
      9100   FORMAT (//,5X,'**********  OUTPUT  **********')
            WRITE (6,9200) PRMSCL (1)
      9200  FORMAT (//,5X,'PROBABILITY OF MISCLASSIFYING AN OBS. FROM POP.1 ='
           1,F7.5)
            WRITE (6,9300) PRMSCL (2)
      9300  FORMAT (//,5X,'PROBABILITY OF MISCLASSIFYING AN OBS. FROM POP.2 ='
```

```
      1,F7.5)
 9400     CONTINUE
          WRITE(6,906)
  906     FORMAT(//,2X,70('$'))
  907    WRITE(6,910)
  910     FORMAT(/,2X,'DO YOU HAVE A CANDIDATE VECTOR TO BE CLASSIFIED')
            READ(5,915)IX
  915     FORMAT(A6)
          IF(IX .NE. IRES) GO TO 999
          WRITE(6,920)
  920     FORMAT(/,2X,'ENTER CANDIDATE VECTOR')
          READ(5,60,END=999) (CANVEC(J),J=1,NL)
          CALL CLSFY (ZOPTCT,KRSIDE,LSIDE)
          WRITE(6,930)
  930     FORMAT(/,2X,'$$$   CANDIDATE VECTOR   $$$')
          WRITE(6,57) (CANVEC(J),J=1,NL)
          WRITE(6,940) KPOP
  940     FORMAT(/,2X,'$$$  CANDIDATE IS CLASSIFIED INTO POPULATION '
      1 ,I2,'  $$$')
            GO TO 907
  900     WRITE (6,905)
  905     FORMAT (//,2X,'ERROR IN GJR***********')
  999  CONTINUE
       END
```

```
      SUBROUTINE CUTOFF(ZBARX,ZBARY,ZCUT,ZOPTCT,LEPSCK)
   COMMON /SEVEN/COST(2),PI(2),PMSCLS(2)
      ZCUT = (ZBARX+ZBARY)/2.
C
C***   INITIALIZE VARIABLES FOR ZCUT OPTIMUM SEARCH
C
    PMSCLS(1) = 1.
    PMSCLS(2) = 1.
   ZOPTCT = ZCUT
   LEPSCK = 0
      RETURN
      END
```

MISSCLASS.OPTZ

```
              SUBROUTINE OPTZ(ZCUT,LEPSCK,ZOPTCT)
              COMMON /SEVEN/COST(2),PI(2),PMSCLS(2)
              COMMON /EIGHT/PRMSCL(2),ALPHA,BETA,EPS,DEL
C**
C*******   CALCULATE EXPECTED COST OF MISCLASSIFICATION
C**
C
              XNWCST = COST(1)*PI(1)*PRMSCL(1) + COST(2)*PI(2)*PRMSCL(2)
              OLDCST = COST(1)*PI(1)*PMSCLS(1) + COST(2)*PI(2)*PMSCLS(2)
           IF(ABS(XNWCST-OLDCST).LT. EPS)GO TO 500
             IF(XNWCST.GT.OLDCST) GO TO 300
             OLDCST = XNWCST
             PMSCLS(1) = PRMSCL(1)
             PMSCLS(2) = PRMSCL(2)
             ZOPTCT = ZCUT
             DEL = ALPHA*DEL
             ZCUT = ZOPTCT + DEL
             GO TO 900
  300        DEL = BETA*DEL
             ZCUT = ZOPTCT + DEL
             GO TO 900
  500        ZCUT = (ZOPTCT + ZCUT)/2.
             LEPSCK = LEPSCK + 1
  900            CONTINUE
              RETURN
              END
```

MISSCLASS.XVEC

```
      SUBROUTINE XVEC1(RNORM1,UNIF)
      COMMON /ONE/XBAR(30),YBAR(30),I
      COMMON /THREE/CMAT(30,30)
      COMMON /FOUR/XVEC(30),BUF(30),ZVEC(30)
      COMMON /FIVE/N
      DO 27 L=1,N,2
      ZVEC(L)=RNORM1(UNIF,RNORM2,0.0,1.0)
      LL=L+1
      ZVEC(LL)=RNORM2
   27 CONTINUE
      DO 121 K=1,N
      SUM=0.0
      DO 111 J=1,N
      SUM=SUM+CMAT(K,J)*ZVEC(J)
  111 CONTINUE
      BUF(K)=SUM
  121 CONTINUE
       IF (I.NE.1) GO TO 150
      DO 131 K=1,N
      XVEC(K)=BUF(K)+XBAR(K)
  131 CONTINUE
          GO TO 900
  150  DO 160 K=1,N
       XVEC(K) = BUF(K) + YBAR(K)
  160  CONTINUE
  900  CONTINUE
      RETURN
      END
```

MISSCLASS.CLSFY

```
      SUBROUTINE CLSFY(ZOPTCT,KRSIDE,LSIDE)
     COMMON /NINE/ CANVEC(30), KPOP,COPY(30,31)
    COMMON /FIVE/NL
    NC = NL + 1
      ZCAN = 0.0
      DO 80  J=1,NL
        ZCAN = COPY(J,NC)*CANVEC(J) + ZCAN
80    CONTINUE
      IF (ZCAN .GE. ZOPTCT) KPOP = KRSIDE
      IF (ZCAN .LT. ZOPTCT) KPOP = LSIDE
      RETURN
      END
```

```
      SUBROUTINE CMAT1
      COMMON /FIVE/N
      COMMON /THREE/CMAT(30,30)
      COMMON /TWO/SIGMA(30,30)
      DO 110 J=1,N
      IF(J.GE.2) GO TO 91
      DO 81 I=1,N
      CMAT(I,1)=SIGMA(I,1)/SQRT(SIGMA(1,1))
   81 CONTINUE
      GO TO 110
   91 DO 105 I=1,N
      IF(J.GE.I+1) GO TO 104
      IF(J.NE.I) GO TO 95
      SUB1=0.0
      L=I-1
      DO 93 K=1,L
      SUB1=SUB1+CMAT(I,K)**2
   93 CONTINUE
      CMAT(I,J)=SQRT(SIGMA(I,J)-SUB1)
      GO TO 105
   95 SUB2=0.0
      L=J-1
      DO 97 K=1,L
      SUB2=SUB2+CMAT(I,K)*CMAT(J,K)
   97 CONTINUE
      CMAT(I,J)=(SIGMA(I,J)-SUB2)/CMAT(J,J)
      GO TO 105
  104 CMAT(I,J)=0.0
  105 CONTINUE
  110 CONTINUE
      RETURN
      END
```

FUNCTION SIMNORM

```
FUNCTION RNORM1(UNIF,RNORM2,U,SIG2)
TPI=6.2831852
A=UNIF(X)
B=UNIF(X)
RNORM1=U+SQRT(-2.0*SIG2*ALOG(A))*COS(TPI*B)
RNORM2=U+SQRT(-2.0*SIG2*ALOG(A))*SIN(TPI*B)
RETURN
END
```

FUNCTION UNIF

```
      FUNCTION UNIF(A)
      DATA IY/96581/
      IY=IY*3125
      IF(IY) 5,6,6
    5 IY=IY+1+34359738367
    6 YFL=IY
      UNIF=YFL*2.0**(-35)
      RETURN
      END
```

```
C     FCLASS IS AN EMPIRICAL SIMULATION TECHNIQUE COMPUTER PROGRAM
C     USING NON-LINEAR DISCRIMINATION
C
C
C***************   VARIABLE LIST   ************************
C
C   ALPHA--ACCELERATION CONSTANT USED IN SEARCH ROUTINE
C   BETA--DECCELERATION CONSTANT USED IN SEARCH ROUTINE
C   BUF--ARRAY USED TO GENERATE OBSERVATIONS (GEN. OBS. WITH MEAN 0.)
C   CANVEC--CANDIDATE VECTOR TO BE CLASSIFIED
C   CLASS(I)-- NUMBER OF GENERATED OBSERVATIONS FROM POP. I CLASSIFIED
C             INTO POPULATION 1
C   CMAT--ARRAY USED IN OBSERVATION GENERATION--C'C=SIGMA
C   COPY--ARRAY USED BY MATHSTAT TO INVERT COVARIANCE MATRICES
C   D--VECTOR OF MEAN DIFFERENCES BETWEEN THE TWO POPULATIONS
C   DEL--STEP LENGTH FOR SEARCH
C   EPS--INTERVAL IN WHICH THE OPTIMUM COST IS FROM SEARCH OPTIMUM
C   F--ARRAY FOR STORAGE OF POPULATION DENSITY FOR OBSERVATIONS
C      F(I,J,K) = ITH POPULATION DENSITY
C      J TH OBSERVATION GENERATED FROM POP. K
C   INDEX--POPULATION IDENTIFIER
C   JD--ARRAY NEEDED FOR MATHPACK CALLS
C   KNT--MAX NUMBER OF ITERATIONS TO BE USED IN SEARCH
C   KPOP--POPULATION IN WHICH CANDIDATE VECTOR IS CLASSFIED
C   LEPSCK--FLAG TO DENOTE END OF SEARCH
C   NK--NUMBER OF OBSERVATIONS GENERATED FROM EACH POPULATION - SET AT 10000
C   NL--NUMBER OF VARIATES COMPOSING AN OBSERVATION
C   NSTART--NUMBER OF STARTS FOR RANDOM NUMBER GENERATOR
C   OLDCST--OPTIMAL EXPECTED COST OF MISCLASSIFICATION
C   PI(I)--PRIOR PROBABILITY THAT AN OBSERVATION COMES FROM POP. I
C   PMSCLS--NON-OPTIMAL PROBABILITY OF MISCLASSIFICATION USED IN SEARCH
C   PRMSCLS(I)--OPTIMAL PROBABILITY OF MISCLASSIFYING AN OBSERVATION FROM POP. I
C   PROD--ARRAY USED TO STORE INTERIM MATRIX PRODUCTS WHEN CALCULATING
C           DENSITY EXPONENT
C   V--IDENTIFIERS USED FOR MATHPACK CALLS (REF. MATHPACK GJR)
C   XBAR--MEAN VECTOR FOR POP.1
C   XSIGIN--INVERSE OF POP.1 COVARIANCE MATRIX
C   XSIGMA--COVARIANCE MATRIX FOR POP1
C   XVEC--GENERATED OBSERVATION VECTOR
C   YBAR--MEAN VECTOR FOR POP.2
C   YSIGIN--INVERSE OF POP.2 COVARIANCE MATRIX
C   YSIGMA--COVARIANCE MATRIX FOR POP.2
C   ZETA--CONSTANT USED TO CLASSIFY OBSERVATIONS
C   ZVEC--ARRAY USED TO GENERATE OBSERVATIONS
C           (VECTOR OF UNIFORM DEVIATES ON INTERVAL 0.,1.)
C
C**************************************************************************
C
C
C*************   SUBROUTINE LIST   ********************
C
C   CLASSIFY--CLASSIFIES CANDIDATE VECTORS -- CLSFY
C   CMAT--COMPUTES CMAT MATRIX
C   INITIALIZE--ASSIGNS VALUES TO ALPHA,BETA,DEL,EPS,ZETA,LEPSCK,OLDCST -- INIT
C   MAIN--READS INPUT,PRINTS OUTPUT,CALCULATES PROBABILITY OF MISCLASSIFICATION
C   OPTIMIZE--OPTIMIZES PROB. OF MISCLASSIFICATION SUCH THAT EXPECTED COSTS
```

```
C              ARE MINIMIZED
C   XVEC--GENERATES OBSERVATION VECTORS
C
C*********************************************************************************
         DIMENSION D(30),JD(30),V(2)
     1,PROD(30,30),F(2,10000,2),CLASS(2)
          COMMON /ONE/XBAR(30),YBAR(30),INDEX
          COMMON /TWO/XSIGMA(30,30),YSIGMA(30,30)
          COMMON /THREE/CMAT(30,30)
          COMMON /FOUR/XVEC(30),BUF(30),ZVEC(30)
        COMMON /FIVE/NL
        COMMON /SEVEN/COST(2),PI(2),PMSCLS(2),PRMSCL(2)
       COMMON /EIGHT/ALPHA,BETA,EPS,OLDCST,OLDZ,LEPSCK,ZETA,DEL
        COMMON /NINE/ CANVEC( 30),  KPOP ,COPY(30,31)
          COMMON /SIX/XSIGIN(30,30),YSIGIN(30,30)
          DATA/IRES/6HYES    /
      EXTERNAL UNIF,RNORM1,GJR,MXMLT
   47    FORMAT(/,2X,'IS COV. MATRIX FOR POP. #1 THE SAME FOR POP.#2?')
   48    FORMAT(//,5X,'**  COVARIANCE MATRIX FOR POP.#2  **')
   49 FORMAT (//,5X,'*****      INPUT      *****')
   50 FORMAT (//,5X,'**  COVARIANCE MATRIX FOR POP#1  **')
   51 FORMAT (//,5X,'**  C  MATRIX  **')
   52 FORMAT (//,5X,'**  MEAN VECTOR FOR POPULATION 1  **')
   53 FORMAT (//,5X,'**  MEAN VECTOR FOR POPULATION 2  **')
   57 FORMAT (/,2X,8(1X,F8.4))
   60 FORMAT ( )
   75     FORMAT(/,2X,'ENTER THE PRIOR PROBABILITY THAT AN OBS. COMES FR
     1OM POP. 1')
   77    FORMAT(/,2X,'ENTER PRIOR PROBABILITY THAT AN OBS. COMES FROM PO
     1P. 2')
   80  FORMAT(/,2X,'**  COST OF MISCLASSIFYING AN OBS. FROM POP. 1 =',
     1F10.5)
   81  FORMAT(/,2X,'**  COST OF MISCLASSIFYING AN OBS. FROM POP. 2 =',
     1F10.5)
   82  FORMAT(/,2X,'**  PRIOR PROB. AN OBS. BELONGS TO POP. 1 =',
     1F10.5)
   83  FORMAT(/,2X,'**  PRIOR PROB. AN OBS. BELONGS TO POP. 2 =',
     1F10.5)
C*****     READ INPUT
C**
C**
       WRITE (6,90)
   90    FORMAT(/,2X,'ENTER NUMBER OF STARTS FOR UNIF')
         READ (5,60,END=999) NSTART
       DO 95 KN=1,NSTART
          X = UNIF(A)
   95   CONTINUE
          WRITE(6,97)
   97     FORMAT(/,2X,'ENTER MAX NUMBER OF ITERATIONS FOR SEARCH')
          READ (5,60,END=999) KNT
      WRITE (6,101)
  101  FORMAT (/,2X,'ENTER DIMENSION OF POPULATIONS')
      READ (5,60,END=999) NL
      WRITE (6,103)
  103  FORMAT (/,2X,'ENTER THE ESTIMATED MEAN VECTOR FOR POPULATION #1')
      READ (5,60,END=999) (XBAR(J),J=1,NL)
      WRITE (6,105)
```

```
  105   FORMAT (/,2X,'ENTER THE ESTIMATED MEAN VECTOR FOR POPULATION #2')
       READ(5,60,END=999) (YBAR(J),J=1,NL)
           WRITE(6,106)
   106     FORMAT(/,2X,'ENTER THE ESTIMATED WITHIN POP#1 COVARIANCE MATRI
     1X')
           READ(5,60,END=999)((XSIGMA(K,J),J=1,NL),K=1,NL)
           WRITE (6,47)
           READ(5,915) IANS
           IF(IANS .EQ. IRES) GO TO 110
       WRITE (6,107)
  107   FORMAT (/,2X,'ENTER THE ESTIMATED WITHIN POP.#2 COVARIANCE MATRIX
     1')
       READ (5,60,END=999) ((YSIGMA(K,J),J=1,NL),K=1,NL)
           GO TO 117
   110     DO 112 K=1,NL
           DO 111 J=1,NL
           YSIGMA(K,J) = XSIGMA(K,J)
   111      CONTINUE
   112      CONTINUE
   117      WRITE(6,118)
  118      FORMAT(/,2X,'ENTER THE ESTIMATED COST OF MISCLASSIFYING AN
     1OBS, FROM POP. 1')
         READ (5,60,END=999) COST(1)
         WRITE (6,119)
  119      FORMAT(/,2X,'ENTER THE ESTIMATED COST OF MISCLASSIFYING AN
     1OBS, FROM POP. 2')
         READ (5,60,END=999) COST(2)
          WRITE (6,75)
         READ (5,60,END=999) PI(1)
         WRITE (6,77)
         READ (5,60,END=999) PI(2)
C
C
C*****      PRINT INPUT FOR ACCURACY CHECK
C
       WRITE (6,49)
       WRITE (6,52)
       WRITE (6,57) (XBAR(J),J=1,NL)
       WRITE (6,53)
       WRITE (6,57) (YBAR (J),J=1,NL)
       WRITE (6,50)
       DO 120 K = 1,NL
       WRITE (6,57)  (XSIGMA(K,J),J=1,NL)
  120      CONTINUE
           WRITE(6,48)
           DO 121 K=1,NL
           WRITE(6,57) (YSIGMA(K,J),J=1,NL)
   121     CONTINUE
   122      CONTINUE
           WRITE(6,80) COST(1)
           WRITE(6,81) COST(2)
           WRITE(6,82) PI(1)
           WRITE(6,83) PI(2)
C
C*****  CALL GJR TO GET DETERMINANTS FOR XSIGMA AND YSIGMA
C
           DO 1100 K=1,NL
```

```
            DO 1090 J=1,NL
            XSIGIN(K,J) = XSIGMA(K,J)
           COPY(K,J)=XSIGMA(K,J)
 1090       CONTINUE
 1100       CONTINUE
            V(1)=2.
            CALL GJR(COPY,30,30,NL,NL,$900,JD,V)
            DETX = EXP(V(2))
           DETX = SIGN(DETX,V(1))
             DO 1150 K=1,NL
             DO 1140 J=1,NL
             YSIGIN(K,J) = YSIGMA(K,J)
          COPY(K,J) = YSIGMA(K,J)
 1140       CONTINUE
 1150       CONTINUE
            V(1) = 2.
            CALL GJR(COPY,30,30,NL,NL,$900,JD,V)
            V(1) = 2.
            DETY = EXP(V(2))
            DETY = SIGN(DETY,V(1))
C
C******** CALL GJR TO GET SIGMA INVERSE
            V(1) = 1.
            CALL GJR(XSIGIN,30,30,NL,NL,$900,JD,V)
            V(1) = 1.
            CALL GJR(YSIGIN,30,30,NL,NL,$900,JD,V)
C
C
C**   INITIALIZE VARIABLES FOR OPT SEARCH
C
           CALL INIT
C
C*******GENERATE OBSERVATIONS
             DO 4000 INDEX=1,2
            CALL CMAT1(INDEX)
             IF (INDEX.EQ.1) JDEX=2
             IF ( INDEX.EQ.2) JDEX=1
             NK = 10000
             DO 4000 K=1,NK
             CALL XVEC1 (RNORM1,UNIF)
             DO 3000 L=1,NL
             D(L) = XVEC(L) -XBAR(L)
 3000         CONTINUE
            CALL MXMLT(D,XSIGIN,PROD,1,NL,NL,1,30)
            CALL MXMLT(PROD,D,CHI,1,NL,1,1,30)
            F(1,K,INDEX) = (DETX**(-.5))*EXP(-CHI/2.)
           DO 3500 M=1,NL
            D(M) = XVEC(M) -YBAR(M)
 3500        CONTINUE
            CALL MXMLT(D,YSIGIN,PROD,1,NL,NL,1,30)
           CALL MXMLT(PROD,D,CHI,1,NL,1,1,30)
            F(2,K,INDEX) = (DETY**(-.5))*EXP(-CHI/2.)
 4000        CONTINUE
C
C***    CLASSIFY THE GENERATED OBSERVATION VIA F-STAT,
C
 4100       DO 5000 INDEX=1,2
```

```
          IF(INDEX .EQ.1) JDEX=2
          IF(INDEX .EQ. 2) JDEX=1
          CLASS(INDEX) = 0.
          IF(INDEX.EQ.2) GO TO 4600
          DO 4500 K=1,NK
         IF(F(INDEX,K,INDEX) .LT. (ZETA*F(JDEX,K,INDEX))) GO TO 4500
         CLASS(INDEX) = CLASS(INDEX) + 1.
 4500     CONTINUE
          GO TO 4900
 4600     DO 4700 K=1,NK
          IF((ZETA*F(2,K,2)).LT.F(1,K,2)) GO TO 4700
          CLASS(2) = CLASS(2) + 1.
 4700     CONTINUE
 4900     CONTINUE
          PRMSCL(INDEX) = (NK-CLASS(INDEX))/NK
 5000     CONTINUE
C
C***   OPT. ZETA
C
          CALL OPTVAL
          IF (LEPSCK.GT.2) GO TO 9000
          KOUNT = KOUNT + 1
         IF ( KOUNT .GT. KNT) GO TO 7500
          GO TO 4100
 7000   CONTINUE
 7500     WRITE(6,7510)
 7510      FORMAT(//,2X,'%%%%%%  KNT EXCEEDED  %%%%%%')
C
C**
C*****      PRINT OUTPUT
C**
C
 9000  CONTINUE
       WRITE (6,9100)
9100   FORMAT (//,5X,'**********  OUTPUT  **********')
       WRITE (6,9200) PRMSCL (1)
9200  FORMAT (//,5X,'PROBABILITY OF MISCLASSIFYING AN OBS. FROM POP.1 ='
     1,F7.5)
       WRITE (6,9300) PRMSCL (2)
9300  FORMAT (//,5X,'PROBABILITY OF MISCLASSIFYING AN OBS. FROM POP.2 ='
     1,F7.5)
 9400     CONTINUE
 907     WRITE(6,910)
 910      FORMAT(/,2X,'DO YOU HAVE A CANDIDATE VECTOR TO BE CLASSIFIED')
            READ(5,915)IX
 915      FORMAT(A6)
          IF(IX .NE. IRES) GO TO 999
          WRITE(6,920)
 920      FORMAT(/,2X,'ENTER CANDIDATE VECTOR')
          READ(5,60,END=999) (CANVEC(J),J=1,NL)
         CALL CLSFY(DETX,DETY,ZETA)
          WRITE(6,930)
 930      FORMAT(/,2X,'$$$   CANDIDATE VECTOR   $$$')
          WRITE(6,57) (CANVEC(J),J=1,NL)
          WRITE(6,940) KPOP
 940      FORMAT(/,2X,'$$$  CANDIDATE IS CLASSIFIED INTO POPULATION '
     1 ,I2,'  $$$')
```

```
            GO TO 907
  900       WRITE (6,905)
  905       FORMAT (//,2X,'ERROR IN GJR***********')
  999  CONTINUE
       END
```

```
      SUBROUTINE INIT
      COMMON /SEVEN/COST(2),PI(2),PMSCLS(2),PRMSCL(2)
      COMMON /EIGHT/ALPHA,BETA,EPS,OLDCST,OLDZ,LEPSCK,ZETA,DEL
      ALPHA = 3.
      BETA = -.5
      DEL = .1
      EPS = 1.
      ZETA = 1.
      LEPSCK = 0
      OLDZ = 1.
     OLDCST = COST(1)*PI(1) + COST(2)*PI(2)
     RETURN
      END
```

```
      SUBROUTINE XVEC1(RNORM1,UNIF)
      COMMON /ONE/XBAR(30),YBAR(30),INDEX
      COMMON /THREE/CMAT(30,30)
      COMMON /FOUR/XVEC(30),BUF(30),ZVEC(30)
      COMMON /FIVE/N
      DO 27 L=1,N,2
      ZVEC(L)=RNORM1(UNIF,RNORM2,0.0,1.0)
      LL=L+1
      ZVEC(LL)=RNORM2
27    CONTINUE
      DO 121 K=1,N
      SUM=0.0
      DO 111 J=1,N
      SUM=SUM+CMAT(K,J)*ZVEC(J)
111   CONTINUE
      BUF(K)=SUM
121   CONTINUE
        IF (INDEX.NE.1) GO TO 150
      DO 131 K=1,N
      XVEC(K)=BUF(K)+XBAR(K)
131   CONTINUE
          GO TO 900
150     DO 160 K=1,N
        XVEC(K) = BUF(K) + YBAR(K)
160     CONTINUE
900     CONTINUE
      RETURN
      END
```

```
      SUBROUTINE CMAT1 (INDEX)
      COMMON /FIVE/NL
      COMMON /THREE/CMAT(30,30)
      COMMON /TWO/XSIGMA(30,30),YSIGMA(30,30)
          DIMENSION SIGMA(30,30)
          IF (INDEX.EQ.2) GO TO 60
          DO 50 K=1,NL
          DO 48 J=1,NL
          SIGMA(K,J) = XSIGMA(K,J)
 48        CONTINUE
 50       CONTINUE
          GO TO 80
 60       DO 70 K=1,NL
          DO 68 J=1,NL
          SIGMA(K,J) = YSIGMA(K,J)
 68       CONTINUE
 70       CONTINUE
 80       DO 110 J=1,NL
      IF(J.GE.2) GO TO 91
      DO 81 I=1,NL
      CMAT(I,1)=SIGMA(I,1)/SQRT(SIGMA(1,1))
 81   CONTINUE
      GO TO 110
 91   DO 105 I=1,NL
      IF(J.GE.I+1) GO TO 104
      IF(J.NE.I) GO TO 95
      SUB1=0.0
      L=I-1
      DO 93 K=1,L
      SUB1=SUB1+CMAT(I,K)**2
 93   CONTINUE
      CMAT(I,J)=SQRT(SIGMA(I,J)-SUB1)
      GO TO 105
 95   SUB2=0.0
      L=J-1
      DO 97 K=1,L
      SUB2=SUB2+CMAT(I,K)*CMAT(J,K)
 97   CONTINUE
      CMAT(I,J)=(SIGMA(I,J)-SUB2)/CMAT(J,J)
      GO TO 105
 104  CMAT(I,J)=0.0
 105  CONTINUE
 110  CONTINUE
      RETURN
      END
```

```
      SUBROUTINE CLSFY(DETX,DETY,ZETA)
      COMMON /ONE/XBAR(30),YBAR(30),INDEX
      COMMON /SIX/XSIGIN(30,30),YSIGIN(30,30)
      COMMON /NINE/CANVEC(30),KPOP,COPY(30,31)
      COMMON /FIVE/NL
      DIMENSION D(30),PROD(30,30)
      DO 300 L=1,NL
        D(L)=CANVEC(L) - XBAR(L)
300   CONTINUE
      CALL MXMLT(D,XSIGIN,PROD,1,NL,NL,1,30)
      CALL MXMLT(PROD,D,CHI,1,NL,1,1,30)
      XCLS = (DETX**(-.5))*EXP(-CHI/2.)
      DO 400 M=1,NL
        D(M) = CANVEC(M) - YBAR(M)
400   CONTINUE
      CALL MXMLT(D,YSIGIN,PROD,1,NL,NL,1,30)
      CALL MXMLT(PROD,D,CHI,1,NL,1,1,30)
      YCLS = (DETY**(-.5))*EXP(-CHI/2.)
      IF(XCLS.LT.(ZETA*YCLS)) GO TO 800
      KPOP = 1
      GO TO 950
800   KPOP = 2
950   CONTINUE
      RETURN
      END
END PRT
@BRK,T
```

FUNCTION.UNIF

```
      FUNCTION UNIF(A)
      DATA IY/96581/
      IY=IY*3125
      IF(IY) 5,6,6
    5 IY=IY+1+34359738367
    6 YFL=IY
      UNIF=YFL*2.0**(-35)
      RETURN
      END
```

FUNCTION SIMNORM

```
FUNCTION RNORM1(UNIF,RNORM2,U,SIG2)
TPI=6.2831852
A=UNIF(X)
B=UNIF(X)
RNORM1=U+SQRT(-2.0*SIG2*ALOG(A))*COS(TPI*B)
RNORM2=U+SQRT(-2.C*SIG2*ALOG(A))*SIN(TPI*B)
RETURN
END
```

APPENDIX III

SQUAD AUTOMATIC WEAPONS OPERATIONAL TEST DATA

Training Subtest Variable X_1

Basic Firing Qualification Scores

Test Participant	Test Candidates A	B	C
1	76	52	80
2	83	67	77
3	85	31	86
4	85	75	86
5	91	48	80
6	70	57	93
7	72	51	84
8	84	67	83
9	49	73	71
10	62	75	66
11	56	63	78
12	46	70	62
13	55	63	71
14	52	71	66
15	36	75	57
16	65	71	82
17	74	62	81
18	85	83	18
19	73	89	63
20	65	68	39
21	81	82	64
22	56	58	67
23	55	80	77
24	65	76	74

Training Subtest Variable X_2

Transition Firing Scores

Test Participant	Test Candidates A	B	C
1	10	0	30
2	10	0	30
3	20	60	20
4	30	0	10
5	40	10	30
6	10	0	20
7	10	30	10
8	30	20	30
9	0	40	10
10	30	50	20
11	20	20	40
12	40	20	50
13	30	60	40
14	0	60	40
15	30	20	20
16	30	50	50
17	10	40	0
18	0	40	30
19	30	50	30
20	0	70	20
21	10	50	50
22	10	30	40
23	0	40	30
24	10	30	0

Training Subtest Variable X_3

Time for Disassembly

Test Participant	Test Candidates A	B	C
1	33	84	55
2	80	40	45
3	34	21	52
4	38	75	90
5	51	45	58
6	70	130	60
7	55	90	101
8	55	81	65
9	49	75	70
10	35	32	53
11	54	122	75
12	50	70	65
13	44	31	66
14	55	33	105
15	32	41	55
16	32	27	62
17	91	62	127
18	39	93	57
19	35	47	75
20	54	70	120
21	30	53	80
22	65	77	150
23	30	39	82
24	42	54	90

Training Subtest Variable X_4

Time for Assembly

Test Participant	Test Candidates A	B	C
1	107	162	110
2	145	136	115
3	60	189	172
4	87	329	120
5	71	115	90
6	210	339	124
7	140	125	91
8	128	157	105
9	65	245	125
10	50	76	105
11	79	410	105
12	105	125	135
13	85	110	100
14	105	304	120
15	44	141	130
16	56	175	131
17	195	122	164
18	82	235	137
19	55	166	135
20	115	304	155
21	50	153	96
22	140	439	301
23	79	212	120
24	85	187	150

Quickfire Subtest Variable X_5

Average Time to 1st Round - 20 m.

Test Participant	Test Candidates A	B	C
1	2.20	1.75	2.29
2	2.42	2.74	2.51
3	2.19	2.45	2.00
4	2.55	2.18	2.79
5	2.39	1.85	2.54
6	2.78	2.72	2.52
7	2.06	3.22	1.81
8	2.28	2.57	2.55
9	2.41	2.86	2.25
10	2.23	2.53	2.58
11	2.35	2.45	2.47
12	1.95	2.04	2.05
13	2.77	2.72	2.24
14	2.61	2.76	2.98
15	1.91	2.39	2.71
16	2.58	2.43	2.93
17	2.11	2.43	2.15
18	2.08	2.53	1.91
19	2.38	2.83	1.60
20	2.33	3.13	2.78
21	2.68	1.96	3.37
22	2.69	2.99	2.67
23	2.36	2.47	2.44
24	2.30	2.41	2.38

Quickfire Subtest Variable X_6

Average Time to 1st Round - 40 m.

Test Participant	Test Candidates A	B	C
1	2.34	2.13	2.19
2	2.79	2.83	1.82
3	2.62	1.91	2.33
4	2.81	4.35	3.01
5	3.01	2.93	2.31
6	2.42	2.65	1.57
7	2.18	3.11	2.16
8	2.47	2.76	2.60
9	2.63	2.46	2.49
10	2.46	3.44	2.19
11	2.05	3.00	3.84
12	1.69	1.90	2.14
13	2.22	2.48	2.36
14	2.54	2.27	2.30
15	2.00	2.53	2.33
16	1.90	2.22	3.88
17	2.02	2.27	1.93
18	5.67	2.62	1.90
19	1.97	3.05	1.82
20	2.59	1.83	2.21
21	2.36	2.74	1.93
22	1.95	3.01	2.45
23	2.45	2.65	2.35
24	2.40	2.60	2.30

Quickfire Subtest Variable X_7

Average Time to 1st Round - 60 m.

Test Participant	Test Candidates A	B	C
1	2.22	2.18	2.20
2	2.58	3.28	1.65
3	2.05	2.20	2.39
4	2.94	3.44	2.41
5	4.26	2.57	2.11
6	2.08	2.25	2.54
7	1.64	2.87	2.11
8	2.22	2.54	2.30
9	2.65	2.71	2.10
10	2.13	2.72	1.81
11	2.01	2.40	2.25
12	2.18	1.70	1.98
13	4.76	5.17	2.20
14	3.33	2.54	2.77
15	2.56	2.34	2.60
16	2.05	2.85	2.58
17	2.22	2.74	2.35
18	1.85	2.47	2.06
19	1.78	2.84	2.05
20	1.01	2.22	4.80
21	1.71	2.23	1.33
22	1.00	4.09	2.41
23	2.39	2.74	2.31
24	2.30	2.65	2.22

Quickfire Subtest Variable X_8

Average Time to 1st Round - 80 m.

Test Participant	Test Candidates A	B	C
1	2.13	1.70	2.01
2	2.51	3.08	1.80
3	2.15	2.31	2.46
4	2.34	1.89	2.82
5	3.00	2.13	3.36
6	2.35	1.84	2.11
7	2.58	1.84	1.74
8	2.34	4.10	2.35
9	2.41	1.87	2.38
10	2.40	2.74	2.24
11	2.85	3.77	2.82
12	1.87	1.85	2.37
13	3.28	3.30	2.28
14	3.60	1.96	2.45
15	2.12	2.31	2.93
16	2.53	2.39	2.55
17	2.78	3.16	2.38
18	1.62	2.33	3.04
19	2.88	3.07	1.99
20	2.71	2.88	2.87
21	3.15	4.23	2.21
22	0.10	1.86	2.31
23	2.55	2.54	2.43
24	2.50	2.49	2.38

Quickfire Subtest Variable X_9

Average Time to 1st Round - Moving Target

Test Participant	Test Candidates A	B	C
1	2.23	1.81	2.68
2	5.60	3.44	3.61
3	3.65	1.90	3.61
4	2.34	2.09	2.02
5	4.08	2.97	6.24
6	3.26	2.82	3.74
7	2.63	3.58	3.29
8	3.03	3.33	2.92
9	3.22	3.44	3.24
10	3.16	3.71	4.57
11	2.49	3.60	4.12
12	2.50	2.86	2.50
13	3.69	2.99	3.00
14	3.50	2.90	2.43
15	5.03	3.53	3.22
16	4.52	4.97	2.92
17	2.15	2.55	3.32
18	2.88	4.30	2.63
19	3.53	3.18	4.72
20	2.57	2.22	0.00
21	5.74	3.13	2.74
22	0.00	0.00	4.68
23	3.41	3.11	3.28
24	3.40	3.10	3.25

Day Defense Subtest Variable X_{10}

Time to First Hit

Test Participant	Test Candidates A	B	C
1	8.5	9.6	10.1
2	9.9	4.2	12.2
3	17.3	4.7	10.0
4	10.9	7.4	10.0
5	8.0	8.3	5.0
6	3.0	7.6	12.6
7	8.4	5.0	10.2
8	13.5	17.5	18.4
9	8.7	6.9	10.5
10	14.9	10.2	11.2
11	8.8	14.7	17.3
12	3.0	15.4	7.3
13	12.2	10.2	7.1
14	12.9	14.7	16.3
15	17.0	16.7	11.6
16	13.6	16.6	20.1
17	15.7	19.5	19.6
18	13.8	20.8	18.5
19	12.0	16.0	19.8
20	18.2	14.5	20.1
21	18.7	29.9	20.0
22	12.0	16.0	14.0
23	20.0	17.4	18.5
24	11.0	16.6	17.0

Day Defense Subtest Variable X_{11}

Time to Change Magazines

Test Participant	Test Candidates A	B	C
1	15	26	40
2	23	47	35
3	19	18	42
4	24	11	31
5	22	20	60
6	15	13	47
7	15	25	27
8	18	13	45
9	15	20	17
10	24	37	71
11	11	19	11
12	16	9	45
13	18	21	79
14	15	5	22
15	41	10	62
16	21	34	36
17	4	15	35
18	15	12	25
19	20	5	32
20	37	7	41
21	33	12	41
22	41	5	24
23	18	11	12
24	18	12	40

Day Defense Subtest Variable X_{11}

Time to Change Magazines

Test Participant	Test Candidates A	B	C
1	15	26	40
2	23	47	35
3	19	18	42
4	24	11	31
5	22	20	60
6	15	13	47
7	15	25	27
8	18	13	45
9	15	20	17
10	24	37	71
11	11	19	11
12	16	9	45
13	18	21	79
14	15	5	22
15	41	10	62
16	21	34	36
17	4	15	35
18	15	12	25
19	20	5	32
20	37	7	41
21	33	12	41
22	41	5	24
23	18	11	12
24	18	12	40

Attack Subtest Variable X_{12}

Time to 1st Round - Sling/100 Rd. Mag.

Test Participant	Test Candidates A	B	C
1	4.64	6.53	12.05
2	9.07	3.90	2.91
3	3.52	3.36	3.72
4	2.65	9.24	4.80
5	15.27	5.63	2.52
6	3.56	1.01	4.06
7	4.87	1.79	4.16
8	9.17	3.39	8.39
9	3.89	4.88	8.37
10	6.78	3.19	4.65
11	3.79	3.90	6.24
12	4.80	5.78	2.60
13	2.35	3.61	3.99
14	2.83	3.07	7.60
15	2.69	5.34	4.90
16	3.00	4.63	5.67
17	14.47	5.45	10.98
18	5.92	4.10	4.57
19	3.83	6.02	14.27
20	4.44	2.82	3.15
21	5.39	4.69	3.37
22	5.17	4.16	10.18
23	3.50	2.70	3.15
24	7.03	12.46	6.29

Attack Subtest Variable X_{13}

Time to 1st Round - Sling/200 Rd. Mag.

Test Participant	Test Candidates A	B	C
1	1.75	7.59	2.16
2	.73	1.26	69.00
3	4.20	5.23	86.00
4	2.13	3.94	11.46
5	61.23	4.96	1.56
6	4.52	3.88	2.27
7	33.71	2.95	4.62
8	6.74	7.11	4.10
9	10.86	5.51	9.67
10	4.41	7.95	3.62
11	3.47	32.97	4.45
12	8.26	10.33	7.36
13	5.36	4.96	4.40
14	5.25	9.86	3.01
15	2.43	8.10	3.45
16	2.91	3.46	4.15
17	4.62	4.17	3.11
18	3.28	3.91	4.80
19	3.59	4.17	3.19
20	3.43	0.19	3.78
21	2.92	3.21	4.09
22	12.05	2.00	4.18
23	3.14	2.84	2.98
24	2.68	5.60	3.58

Attack Subtest Variance X_{14}

Time to 1st Round - Shoulder/100 Rd. Mag.

Test Participant	Test Candidates A	B	C
1	6.68	8.08	28.10
2	5.25	0.07	10.13
3	12.13	10.19	4.46
4	4.86	17.28	14.47
5	70.68	4.09	7.77
6	1.90	7.66	2.20
7	8.75	4.94	5.23
8	13.86	6.19	10.63
9	19.02	12.38	7.83
10	5.92	10.91	7.13
11	4.59	12.03	4.05
12	21.12	21.90	3.51
13	3.04	10.48	4.49
14	3.52	16.43	4.47
15	3.93	5.52	11.59
16	4.42	9.00	5.47
17	12.09	5.57	4.88
18	25.49	6.75	6.71
19	11.72	8.66	4.19
20	4.66	3.21	9.08
21	9.68	3.96	8.46
22	9.58	3.63	4.77
23	7.29	11.28	3.49
24	9.09	3.63	5.35

Attack Subtest Variable X_{15}

Time to 1st Round - Shoulder/200 Rd. Mag.

Test Participant	Test Candidates A	B	C
1	14.02	8.28	18.01
2	4.43	1.91	6.99
3	6.33	7.70	3.87
4	7.86	5.53	21.40
5	9.08	4.06	3.49
6	49.68	4.93	2.27
7	4.31	5.60	9.79
8	5.14	5.40	3.90
9	34.10	8.64	9.40
10	7.42	5.76	7.42
11	14.06	15.13	22.43
12	15.69	6.13	5.40
13	3.62	9.47	4.46
14	15.85	13.88	3.78
15	5.04	10.50	9.62
16	0.89	17.20	9.26
17	27.37	5.77	4.24
18	3.26	4.97	5.51
19	3.86	12.98	12.80
20	3.41	8.91	10.90
21	3.22	6.45	6.93
22	2.84	4.90	5.10
23	3.99	12.58	5.74
24	6.72	4.60	4.25

Attack Subtest Variable X_{16}

Magazine Change Times - 200→100 Rd.

Test Participant	Test Candidates A	B	C
1	57	4	22
2	20	7	42
3	55	9	25
4	22	5	46
5	52	11	44
6	18	12	21
7	9	7	33
8	20	13	66
9	31	10	31
10	18	7	39
11	12	9	43
12	76	13	24
13	35	5	31
14	31	10	66
15	38	5	37
16	10	6	35
17	13	6	25
18	12	11	65
19	12	8	18
20	17	8	126
21	20	4	43
22	21	5	164
23	21	5	99
24	16	7	72

Attack Subtest Variable X_{17}

Magazine Change Time - 100→200 Rd.

Test Participant	Test Candidates A	B	C
1	10	12	33
2	16	8	48
3	22	6	38
4	25	7	45
5	13	8	39
6	27	7	34
7	7	5	27
8	19	4	145
9	34	3	22
10	16	4	170
11	33	5	33
12	50	6	25
13	13	4	43
14	27	6	18
15	36	7	38
16	26	8	24
17	11	6	26
18	32	3	27
19	9	6	12
20	24	6	20
21	8	5	23
22	26	6	22
23	26	4	22
24	70	9	17

Attack Subtest Variable X_{18}

Barrel Change Time

Test Participant	Test Candidates A	B	C
1	13	13	38
2	12	26	31
3	10	8	57
4	41	11	50
5	23	27	40
6	30	20	121
7	15	13	61
8	14	21	42
9	11	33	94
10	12	16	99
11	30	26	75
12	23	36	35
13	12	24	22
14	20	8	50
15	61	12	104
16	116	18	32
17	28	17	28
18	20	23	26
19	26	12	26
20	13	38	35
21	16	8	41
22	17	26	82
23	27	17	46
24	12	14	14

Day Defense Subtest Variable X_{19}

% Targets Engaged

Test Participant	Test Candidates A	B	C
1	87.5	62.5	62.5
2	62.5	62.5	7.50
3	75.0	75.0	62.5
4	100.0	87.5	62.5
5	75.0	75.0	37.5
6	37.5	75.0	62.5
7	87.5	62.5	75.0
8	100.0	87.5	100.0
9	62.5	62.5	87.5
10	100.0	87.5	75.0
11	75.0	100.0	75.0
12	62.5	100.0	62.5
13	100.0	75.0	50.0
14	87.5	100.0	87.5
15	100.0	100.0	50.0
16	87.5	100.0	75.0
17	100.0	100.0	100.0
18	100.0	100.0	100.0
19	100.0	100.0	87.5
20	100.0	100.0	87.5
21	100.0	100.0	87.5
22	100.0	100.0	100.0
23	87.5	100.0	87.5
24	100.0	100.0	87.5

REFERENCES

1. Afifi, A. A. and Azen, S. P., Statistical Analysis, A Computer Oriented Approach, Academic Press, New York, 1972.

2. ALMC, A Course of Instruction in Risk Analysis, ALM-3315-H, U.S. Army Logistics Management Center, Fort Lee, Virginia.

3. AMSAA, Compendium on Risk Analysis Techniques, AMSAA SP-4, Army Materiel Systems Analysis Agency, Aberdeen Proving Ground, Maryland.

4. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New York, 1958.

5. Andrews, D. F., "A Note of the Selection of Data Transformations," Biometrika, Vol. 58, No. 2, pp. 249-254.

6. Andrews, D. F., Gnanadesikan, R., and Warner, J. L., "Transformations of Multivariate Data," Biometrics, Vol. 27, pp. 825-840.

7. Andrews, D. F., Gnanadesika, R., and Warner, J. L., "Methods for Assessing Multivariate Normality," Multivariate Analysis III, Paruchuri R. Krishnaian, editor, Academic Press, Inc., New York, 1973.

8. AR 10-4, dated 30 December 1974, Organization and Functions of U.S. Army Operational Test and Evaluation Agency, Headquarters, Department of the Army, Washington, D.C.

9. Bolch, B. W. and Huang, C. J., Multivariate Statistical Methods for Business and Economics, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1974.

10. Box, G. E. P. and Cox, D. R., "An Analysis of Transformations," J. R. Statist. Soc. B, Vol. 26, pp. 211-252, 1964.

11. Broffitt, J. and Williams, J., "Minimum Variance Estimators for Misclassification Probabilities in Discriminant Analysis," Journal of Multivariate Analysis, Vol. 3, pp. 311-327.

12. Cochran, W. G., "On the Performance of the Linear Discriminant Function," Technometrics, Vol. 6, pp. 179-190.

13. Cochran, W. G. and Bliss, C. I., "Discriminant Functions with Covariance," Ann. Math. Stat., Vol. 19, pp. 151-176.

14. Cooley, W. and Lohnes, P., Multivariate Procedures for the Behavioral Sciences, John Wiley and Sons, New York, 1962.

15. Dempster, A. P., Elements of Continuous Multivariate Analysis, Addison-Wesley Publishing Co., Reading, Mass., 1969.

16. Dixon, W. J., BMD Biomedical Computer Programs, 3rd Edition, University of California Press, Berkeley, 1974.

17. Dolby, J. L., "On a Quick Method of Choosing a Transformation," Technometrics, Vol. 5, pp. 317-325.

18. Fisher, R. A., "The Statistical Utilization of Multiple Measurements," Annals of Eugenics, Vol. VIII, Part IV, pp. 376-386, 1938.

19. Fisher, R. A., "The Precision of Discriminant Functions," Annals of Eugenics, Vol. X, Part IV, pp. 122-429, 1940.

20. Fishman, G. S., Concepts and Methods in Discrete Event Digital Simulation, John Wiley and Sons, New York, 1973.

21. Fix, E. and Hodges, J. L., "Discriminatory Analysis," Project Report 12-49-004, Numbers 4 and 11, U.S. Air Force School of Aviation Medicine, Randolph Field, San Antonio, Texas, 1951.

22. Frank, R. E., Massy, W. F., and Morrison, D. G., "Bias in Multiple Discriminant Analysis," Journal of Marketing Research, Vol. 2, pp. 250-258.

23. Gnanadesikan, R. and Wilk, M. B., "Data Analytic Methods in Multivariate Statistical Analysis," Multivariate Analysis II, P. R. Krishnaian, editor, pp. 593-638, Academic Press, New York, 1969.

24. John, S., "Errors in Discrimination," Ann. Math Statis., Vol. 32, pp. 1125-1144.

25. Kabe, D. G. and Gupta, R. P., Multivariate Statistical Inference, North-Holland Publishing Co., Amsterdam, 1973.

26. Kendall, M. G., A Course in Multivariate Analysis, Charles Griffin Co., Ltd., London, 1963.

27. Lachenbruch, P. A., "An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis," Biometrics, Vol. 23, pp. 639-645.

28. Lachenbruch, P. A., "On Expected Probabilities of Misclassification in Discriminant Analysis, Necessary Sample Size, and a Relation with the Multiple Correlation Coefficient," Biometrics, Vol. 24, pp. 823-834.

29. Lachenbruch, P. A. and Mickey, M. R., "Estimation of Error Rates in Discriminant Analysis," Technometrics, Vol. 10, pp. 1-11.

30. Lilliefors, H., "On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown," Journal of the American Statistical Association, Vol. 62, pp. 399-402.

31. Mahalanobis, P. C., "On the Generalized Distance in Statistics," Proceedings of the National Institute of Science, India, Vol. 12, pp. 49-55.

32. Malkovich, J. F. and Afifi, A. A., "On Tests for Multivariate Normality," Journal of the American Statistical Association, Vol. 68, No. 341, pp. 176-179.

33. Massy, W. F., "On Methods: Discriminant Analysis of Audience Characteristics," Journal of Advertising Research, Vol. 5, pp. 39-48.

34. Morrison, D. F., Multivariate Statistical Methods, McGraw-Hill Book Co., New York, 1967.

35. McQueen, J. B., "Some Methods for Classification and Analysis of Multivariate Observations," Proc. Symp. Math. Statist. and Probability, 5th, Berkeley, Vol. 1, pp. 281-297, University of California Press, Berkeley.

36. Okamoto, M., "An Asymptotic Expansion for the Distribution of the Linear Discriminant Function," Ann. Math. Statist., Vol. 34, pp. 1286-1301; Correction: Ann. Math. Statist., Vol. 39, p. 1358.

37. Operations Research Associates, "Techniques and Guidelines," Operational Test Methodology Guide, Vol. II, January 1974.

38. Overall, J. and Keeet, J., Applied Multivariate Analysis, McGraw-Hill Book Co., New York, 1971.

39. Press, S. J., Applied Multivariate Analysis, Holt, Rinehart and Winston, Inc., New York, 1972.

40. Rao, C. R., "Some Characterizations of the Multivariate Normal Distribution," Multivariate Analysis II, P. R. Krishnaian, editor, Academic Press, Inc., New York, 1969.

41. Sedransk, N. and Okamoro, M., "Estimation of the Probabilities of Misclassification for a Linear Discriminant Function in the Univariate Normal Case," Ann. Inst. Statist. Math., Vol. 23, pp. 419-436.

42. Schuer, E. and Stoller, D. S., "On the Generation of Normal Random Vectors," Technometrics, Vol. 4, pp. 278-281.

43. Sitgreaves, R., "On the Distribution of Two Random Matrices Used in Classification Procedures," Ann. Math. Stat., Vol. 23, p. 263.

44. Sorum, M., "Estimating the Expected Probability of Misclassification for a Rule Based on the Linear Discriminant Function: Univariate Normal Case," Technometrics, Vol. 15, pp. 327-339.

45. Sorum, M., "Three Probabilities of Misclassification," Technometrics, Vol. 14, pp. 309-316.

46. Sorum, M., "Estimating the Expected and the Optimal Probabilities of Misclassification," Technometrics, Vol. 14, pp. 935-943.

47. Sorum, M., "Estimating the Conditional Probability of Misclassification," Technometrics, Vol. 13, pp. 333-343.

48. Tatsuoka, M. M., Multivariate Analysis, John Wiley and Sons, Inc., New York, 1971.

49. Tukey, J. W., "On the Comparative Anatomy of Transformations," Ann. Math. Statist., Vol. 28, pp. 602-632.

50. USAF Academy Risk Analysis Study Team, "Final Report of the USAF Academy Risk Analysis Study Team," Wright-Patterson AFB, Ohio, August 1, 1971.

51. Van de Geer, J. P., Introduction to Multivariate Analysis, W. H. Freeman and Co., San Francisco, 1971.

52. Wald, A., "On a Statistical Problem in the Classification of an Individual into One of Two Groups," Ann. Math. Stat., Vol. 15, p. 145.

53. Welch, B., "Note on Discriminant Functions," Biometrika, Vol. 31, p. 218.